APR 19 1921

GENERAL LIBRARY
APR 19 1921
UNIV. OF MICH.

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS:

HAROLD ORDWAY RUGG, *Chairman.*
Teachers College, Columbia University.
Lincoln School of Teachers College.

JAMES CARLETON BELL,
Brooklyn Training School for Teachers.

FRANK NUGENT FREEMAN,
University of Chicago.

ARTHUR IRVING GATES,
Teachers College, Columbia University.

VIVIAN ALLEN CHARLES HENMON,
University of Wisconsin.

RUDOLF PINTNER,
Ohio State University.

LEWIS MADISON TERMAN,
Leland Stanford University.

EDWARD LEE THORNDIKE,
Teachers College, Columbia University.

LAURA ZIRBES, *Assistant Editor.*
Lincoln School of Teachers College.

VOL. XII.
No. 4.

APRIL, 1921

$4.00 a year.
60c. per copy.

CONTENTS

THE PSYCHOLOGY OF DRILL IN ARITHMETIC: THE AMOUNT OF PRACTICE. *E. L. Thorndike* 183

INTELLIGENCE AND ITS MEASUREMENT: A SYMPOSIUM.

VIII. *By V. A. C. Henmon* 195
IX. *By Joseph Peterson* 198
X. *By L. L. Thurstone* 202
XI. *By Herbert Woodrow* 207
XII. *By W. F. Dearborn* 211
XIII. *By M. E. Haggerty* 212

THE RELATIONSHIP BETWEEN EYE-PERCEPTION AND VOICE RESPONSE IN READING. *G. T. Buswell* 217

PROPHECY OF LEARNING PROGRESS BY BETA. *Garry C. Myers* 228

DEPARTMENT FOR DISCUSSION OF RESEARCH PROBLEMS 232

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 236

Published Monthly Except June to August by
WARWICK and YORK, Inc., Baltimore, Md.

Entered at the Baltimore, (Md.) Postoffice as second-class mail. January 24, 1921.

The Journal of Educational Psychology

is published monthly, except June, July and August, by Warwick & York, Inc., 10 E. Center street, Baltimore, Md. The issues of 1921 make Volume XII.

Manuscripts for publication, books or other materials for review and news items should be addressed to Harold Ordway Rugg, Chairman of the Editorial Board, 59 Edgecliff Terrace, Yonkers, N. Y.

The price of the Journal is $4 a year in advance. Foreign postage is 30 cents extra. Single issues are 60 cents each, and less than a full year's subscription is charged at single-issue rate.

Subscribers should notify publishers of changes in address at least four weeks before publication of the issue with which change takes effect.

Educational Psychology Monographs

Studies in the field of educational psychology which are of too great length to be printed in the Journal are issued in book form in the Educational Psychology Monograph Series. These monographs are of varying size and price, and have no definite date for publication. The works listed below have been published, and may be obtained from Warwick & York, Inc., Baltimore, Md.

1. **Moto-Sensory Development.** George V. N. Dearborn. vi, 215 pages, frontispiece. $1.80.
2. **Mental Fatigue.** Max Offner; translated by G. M. Whipple. viii, 133 pages. $1.50.
3. **Backward and Feeble-Minded Children.** Edmund B. Huey. 221 pages, illus. (Out of print.)
4. **When Should a Child Begin School.** W. H. Winch. 108 pages. $1.40.
5. **Spelling Efficiency in Relation to Age, Grade and Sex, and the Question of Transfer.** J. E. Wallace Wallin. vi, 91 pages. $1.40.
6. **Relative Efficiency of Phonetic Alphabets.** Guy Montrose Whipple. 60 pages. 60 cents.
7. **Experimental Studies of Mental Defectives.** J. E. Wallace Wallin. 156 pages, illus. $1.60.
8. **Variations in the Grades of High School Pupils.** Clarence Truman Gray. 120 pages. $1.50.
9. **How I Kept My Baby Well.** Anna G. Noyes. 193 pages and 60 illustrations. $1.80.
10. **The Marking System in Theory and Practice.** I. E. Finkelstein. 88 pages. $1.40.
11. **Inductive Versus Deductive Methods in Teaching.** W. H. Winch. 160 pages. $1.80.
12. **Children's Perceptions.** W. H. Winch. 5-color frontispiece 240 pages. $2.00.
13. **The Psychological Methods of Testing Intelligence.** W. Stern; translated by G. M. Whipple. 160 pages. $1.60.
14. **The Mentality of the Criminal Woman.** Jean Weidensall. 350 pages, illus. $2.50.
15. **The General Value of Visual Sense Training in Children.** Chang Ping Wang. 100 pages. $1.40.
16. **The Doctrine of Formal Discipline in the Light of Experimental Investigation.** Nellie P. Hewins. 120 pages. $1.50.
17. **The Experimental Determination of Mental Discipline in School Studies.** Harold O. Rugg. 150 pages. $1.75.
18. **The Stanford Revision and Extension of the Binet-Simon Intelligence Scale.** Lewis M. Terman and others. 184 pages. $1.80.
19. **The Psychology and Pedagogy of Anger.** Roy Franklin Richardson. 120 pages. $1.40.
20. **The Picture Completion Test.** Rudolf Pintner and Margaret M. Anderson. 112 pages, illus. $1.50.
21. **Deficiency and Delinquency: An Interpretation of Mental Testing.** James Burt Miner. 380 pages. $2.50.
22. **The Influence of Age and Experience on Correlations Concerned With Mental Testing.** Edward Safford Jones. 98 pages. $1.40.
23. **Periodicity With Reference to Mental and Physical Ability.** A. G. Peaks. 120 pages. $1.40. (In process.)
24. **Graphology and the Psychology of Handwriting.** June E. Downey. 180 pages. $1.80.

Warwick & York, Inc., Baltimore, Md.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Vol. XII APRIL, 1921 No. 4

THE PSYCHOLOGY OF DRILL IN ARITHMETIC: THE AMOUNT OF PRACTICE*

E. L. THORNDIKE,
Teachers College, Columbia University.

THE AMOUNT OF PRACTICE.

It will be instructive if the reader will perform the following experiment as an introduction to the discussion of this article before reading any of the discussion.

Suppose that a pupil does all the work, oral and written, computation and problem solving, presented for grades I to VI inclusive (that is, in the first two books of a three-book series) in the average text-book now used in the elementary school. How many times will he have exercised each of the various bonds involved in the four operations with integers. That is, how many times will he have thought "1 and 1 are 2," "1 and 2 are 3," etc., etc. Every case of the action of each bond is to be counted.

Since estimating for the entire series is too long a task, it will be sufficient to use eight or ten from each, say:

3 + 2	13, 23, etc., + 2	7 + 2	17, 27, etc., + 2
" 3	" 3	" 3	" 3
" 4	" 4	" 4	" 4
" 5	" 5	" 5	" 5
" 6	" 6	" 6	" 6
" 7	" 7	" 7	" 7
" 8	" 8	" 8	" 8
" 9	" 9	" 9	" 9

*Copyright, 1921, by E. L. Thorndike.

3 — 3	7 — 7	9 × 7	63 ÷ 9
4 "	8 "	7 × 9	64 "
5 "	9 "		65 "
6 "	10 "		66 "
7 "	11 "	8 × 6	67 "
8 "	12 "	6 × 8	68 "
9 "	13 "		69 "
10 "	14 "		70 "
11 "	15 "		71 "
12 "	16 "		

TABLE I.

Estimates of the amount of practice provided in Books I and II of the average three-book text in Arithmetic; by 50 experienced teachers.

Arithmetical fact.	Estimate. Lowest.	Median.	Highest.	Range required to include half of the estimates.
3 or 13 or 23, etc., + 2	25	1,500	1,000,000	800—5,000
" 3	24	1,450	80,000	475—5,000
" 4	23	1,150	50,000	750—5,000
" 5	22	1,400	44,000	700—5,000
" 6	21	1,350	41,000	700—4,500
" 7	21	1,500	37,000	600—4,000
" 8	20	1,400	33,000	550—4,100
" 9	20	1,150	28,000	650—4,500
7 or 17 or 27, etc., + 2	20	1,250	2,000,000	600—5,000
" 3	19	1,100	1,000,000	650—4,900
" 4	18	1,000	80,000	650—4,900
" 5	17	1,300	80,000	650—4,400
" 6	16	1,100	29,000	650—4,500
" 7	15	1,100	25,000	500—4,500
" 8	13	1,100	21,000	650—3,800
" 9	10	1,275	17,000	500—4,000
3 — 3	25	1,000	100,000	500—4,000
4 "	20	1,050	500,000	525—3,000
5 "	20	1,100	2,500,000	650—4,200
6 "	10	1,050	21,000	650—3,250
7 "	22	1,100	15,000	550—3,050
8 "	21	1,075	15,000	650—3,000
9 "	21	1,000	15,000	700—2,600
10 "	20	1,000	20,000	600—2,500
11 "	20	1,000	15,000	465—2,550
12 "	18	1,000	15,000	650—2,100
7 — 7	10	1,000	18,000	425—3,000
8 "	15	1,000	18,000	413—3,100
9 "	15	950	18,000	550—3,000
10 "	15	950	18,000	600—3,950
11 "	10	900	18,000	550—3,000
12 "	10	925	18,000	525—3,100
13 "	10	900	18,000	500—2,600
14 "	10	900	18,000	500—3,100
15 "	10	925	18,000	500—3,000
16 "	10	875	18,000	500—2,500

TABLE I—(Continued.)

Arithmetical fact.	Estimate. Lowest.	Median.	Highest.	Range required to include half of the estimates.
9 × 7	10	700	20,000	500—2,000
7 × 9	10	700	20,000	500—1,750
8 × 6	10	750	20,000	500—2,500
6 × 8	9	700	20,000	500—2,500
63 ÷ 9	9	500	4,500	300—2,500
64 "	9	200	4,000	100— 700
65 "	8	200	4,000	100— 600
66 "	7	200	4,000	100— 550
67 "	7	200	4,000	75— 450
68 "	6	200	4,000	87— 575
69 "	6	200	4,000	87— 450
70 "	5	200	4,000	75— 575
71 "	5	200	4,000	75— 700

Having made his estimates the reader should compare them first with similar estimates made by experienced teachers (shown in Table I), and then with the results of actual counts for representative text-books in arithmetic (shown in Tables II to VII).

It will be observed in Table I that even experienced teachers vary enormously in their estimates of the amount of practice given by an average text-book in arithmetic, and that most of them are in serious error by over-estimating the amount of practice. In general it is the fact that we use text-books in arithmetic with very vague and erroneous ideas of what is in them, and think they give much more practice than they do give.

The authors of the text-books as a rule also probably had only very vague and erroneous ideas of what was in them. If they had known, they would almost certainly have revised their books. Surely no author would intentionally provide nearly four times as much practice on $2 + 2$ as on $8 + 8$, or eight times as much practice on 2×2 as on 9×8, or eleven times as much practice on $2 - 2$ as on $17 - 8$, or over forty times as much practice on $2 \div 2$ as on $75 \div 8$ and $75 \div 9$, both together. Surely no author would have provided intentionally only twenty to thirty occurrences each of $16 - 7$, $16 - 8$, $16 - 9$, $17 - 8$, $17 - 9$, and $18 - 9$ for the entire course through grade VI; or have left the practice on $60 \div 7$, $60 \div 8$, $60 \div 9$, $61 \div 7$, $61 \div 8$, $61 \div 9$, and the like to occur only about once a year!

Tables II to VII show that even gifted authors make instruments for instruction in arithmetic which contain much less practice on certain elementary facts than teachers suppose; and which contain relatively much more practice on the more easily learned facts than on those which are harder to learn.

How much practice should be given in arithmetic? How should it be divided amongst the different bonds to be formed? Below a certain amount there is waste because the pupil will need more time to detect and correct his errors than would have been required to give him mastery. Above a certain amount there is waste because of unproductive over-learning. If 668 is just enough for 2×2, 82 is not enough for 9×8. If 82 is just enough for 9×8, 668 is too much for 2×2.

TABLE II.

Amount of Practice: Addition Bonds in a Recent Text-book (A) of Excellent Reputebooks I and II. All save certain supplementary material which does not appear to be considered a part of the normal work.

The Table reads: "$2 + 2$ was used 226 times, $12 + 2$ was used 74 times, $22 + 2$, $32 + 2$, $42 + 2$ and so on were used 50 times."

	2	3	4	5	6	7	8	9	Total.
2	226	154	162	150	97	87	66	45	
12	74	53	76	46	51	37	36	33	
22, etc.	50	60	68	63	42	50	38	26	
3	216	141	127	89	82	54	58	40	
13	43	43	60	70	52	30	22	18	
23, etc.	15	30	51	50	42	32	29	30	
7	85	90	103	103	84	81	61	47	
17	35	25	42	32	35	21	29	16	
27, etc.	30	23	32	29	24	23	25	28	
8	185	112	146	99	75	71	73	61	
18	28	35	52	46	28	29	24	14	
28, etc.	53	36	34	38	23	36	27	27	
9	104	81	112	96	63	74	58	57	
19	13	11	31	38	25	14	22	11	
29, etc.	19	17	27	20	32	32	19	18	
2, 12, 22, etc., +	350	277	306	260	190	174	140	104	1,801
3, 13, 23, etc.	274	214	230	209	176	116	109	88	1,406
7, 17, 27, etc.	148	138	187	164	141	125	115	91	1,109
8, 18, 28, etc.	266	183	232	185	126	136	124	102	1,354
9, 19, 29, etc.	136	109	170	154	120	120	99	86	994
Totals..	1,164	921	1,125	972	753	671	687	471	

It is possible to find the answers to these questions for the pupil of median ability (or any stated ability) by suitable experiments. The amount of practice will of course vary according to the ability of the pupil. It will also vary according to the interest aroused in him and the satisfaction he feels in progress and mastery. It will also vary according to the amount of practice of other related bonds; $7 + 7 = 14$ and $60 \div 7 = 8$ and 4 remainder will help the formation of $7 + 8 = 15$ and $61 \div 7 = 8$ and 5 remainder. It will also of course vary with the general difficulty of the bond, $17 - 8 = 9$ being under ordinary conditions of teaching, harder to form than $7 - 2 = 5$.

Until suitable experiments are at hand we may hazard estimates for the fundamental bonds as follows, assuming that by the end of grade VI a strength of 199 correct out of 200 is to be had, and that the teaching is by an intelligent person working in accord with psychological principles as to both ability and interest.

For one of the easier bonds, most facilitated by other bonds (such as $2 \times 5 = 10$, or $10 - 2 = 8$, or the double bond $7 =$ two threes and 1 remainder) in the case of the median or average pupil, twelve practices in the week of first learning, supported by twenty-five practices during the two months following, and maintained by thirty practices well spread over the later periods. For the more gifted pupils, lesser amounts down to six, twelve and fifteen may suffice. For the less gifted pupils, more may be required up to thirty, fifty and a hundred. It is to be doubted, however, whether pupils requiring so much sheer repetition of these easy bonds as this should be taught arithmetic beyond a few matters of practical necessity.

For bonds of ordinary difficulty, with average facilitation from other bonds (such as $11 - 3$, 4×7, or $48 \div 8 = 6$) in the case of the median or average pupil, twenty practices in the week of first learning, supported by thirty during the two months following, and maintained by fifty practices well spread over the later periods. Gifted pupils may gain and keep mastery with twelve, fifteen and twenty practices respectively. Pupils dull at arithmetic may need up to twenty, sixty and two hundred. Here, again, it is to be doubted whether a pupil for whom arithmetical facts, well-taught and made interesting, are so hard to acquire as this, should learn many of them.

TABLE III.

Amount of Practice: Subtraction Bonds in a Recent Text-book (A) of Excellent Repute: Books I and II. All save certain supplementary material, which does not appear to be considered a part of the normal work.

Frequencies of subtractions of 1 from 1, 1 or 2 from 2, 2 or 3 from 3, etc.

Minuends	Subtrahends. 1	2	3	4	5	6	7	8	9
1	372								
2	214	311							
3	136	149	189						
4	146	142	103	205					
5	171	91	92	164	136				
6	80	59	69	71	81	192			
7	106	57	55	67	59	156	80		
8	73	50	50	75	50	62	48	152	
9	71	75	54	74	48	55	55	124	133
10	261	84	63	100	193	83	57	124	91
11		48	31	50	36	41	32	46	35
12			48	77	57	51	35	80	30
13				35	22	40	29	35	28
14					25	37	36	49	32
15						33	19	48	20
16							16	36	26
17								27	20
18									19
Totals, Exluding 1 — 1, 2 — 2, etc.	1,258	755	565	613	571	558	327	569	301

TABLE IV.

Frequencies of subtractions not included in Table III.

These are cases where the pupil would, by reason of his stage of advancement, probably operate 35 — 30, 46 — 46, etc., each as one bond.

Minuends	Subtrahends. 1	2	3	4	5	6	7	8	9	10
	11	12	13	14	15	16	17	18	19	20
	21 etc.	22 etc.	23 etc.	24 etc.	25 etc.	26 etc.	27 etc.	28 etc.	29 etc.	etc.
10, 20, 30, 40, etc.	11	29	16	52	32	51	7	30	22	60
11, 21, 31, 41, etc.	42	14	22	32	12	26	19	52	17	10
12, 22, 32, 42, etc.	47	97	5	13	9	21	11	24	19	17
13, 23, 33, 43, etc.	7	40	7	14	15	13	19	19	22	3
14, 24, 34, 44, etc.	8	28	14	58	13	16	14	26	19	7
15, 25, 35, 45, etc.	21	28	29	54	51	15	21	12	24	8
16, 26, 36, 46, etc.	5	18	12	27	35	69	13	17	19	2
17, 27, 37, 47, etc.	5	9	12	40	32	54	24	12	12	1
18, 28, 38, 48, etc.	2	16	10	23	22	36	18	47	16	0
19, 29, 39, etc.	5	7	7	10	13	28	14	23	16	0
Totals	153	276	134	323	234	329	160	261	186	117

TABLE V.

Amount of Practice: Multiplication Bonds in Another Recent Text-book (B) of Excellent Repute. Books I and II.

Multipliers	Multiplicands. 0	1	2	3	4	5	6	7	8	9	Totals.
1.....	299	534	472	271	310	293	261	178	195	99	2,912
2.....	350	644	668	480	458	377	332	238	239	155	3,941
3.....	280	487	509	388	318	302	247	199	227	152	3,109
4.....	186	375	398	242	203	265	197	163	159	93	2,281
5.....	268	359	393	234	263	243	217	192	197	114	2,480
6.....	180	284	265	199	196	191	148	169	165	106	1,923
7.....	135	283	277	176	187	158	155	121	145	118	1,755
8.....	137	272	292	175	192	164	158	157	126	126	1,799
9.....	71	173	140	122	97	102	101	100	82	110	1,098
Totals	1,906	3,411	3,414	2,287	2,224	2,095	1,836	1,517	1,535	1,073	

TABLE VI.

Amount of Practice: Divisions without Remainder in Text-book (B), Parts I and II.

Integral Multiples of 2 to 9 in sequence, i. e., $4 \div 2$ occurred 397 times; $6 \div 2$ occurred 256 times; $6 \div 3$, 224 times; $9 \div 3$, 124 times.

Divisors. 2	3	4	5	6	7	8	9	Totals.
397	224	250	130	93	44	98	23	1,259
256	124	152	79	28	43	61	25	768
318	123	130	65	50	19	39	19	763
258	98	86	105	25	24	34	20	650
198	49	76	27	22	30	33	16	451
77	54	36	31	28	27	16	9	278
180	91	50	38	17	13	22	16	427
69	46	37	24	12	17	16	15	236
Totals..... 1,753	809	817	499	275	217	319	142	

For bonds of greater difficulty, less facilitated by other bonds (such as $17 - 9$, 8×7, or $12\frac{1}{2}\%$ of $= \frac{1}{8}$ of), the practice may be from ten to a hundred percent more than the above.

UNDER-LEARNING AND OVER-LEARNING.

Accepting the above provisional estimates as reasonable, we may consider the harm done by giving less and by giving more than these reasonable amounts. Giving less is indefensible. The pupil's time is wasted in excessive checking to find his errors. He is in danger of being practiced in error. His attention is diverted from the learning of new facts and processes by the necessity of thinking out these supposedly mastered facts. All new bonds are harder to learn than they should be because the bonds which should facilitate them are not strong enough to do so. Giving more does harm to

TABLE VII.

Division Bonds, with and without Remainders, Book B.
All work through grade VI, except estimates of quotient figures in long division.

Dividend	2		3			4				5					6						
Divisor	1	2	1	2	3	1	2	3	4	1	2	3	4	5	1	2	3	4	5	6	
No. of occurrences	41	386	27	189	240	26	397	66	185	23	136	43	53	135	21	256	224	68	43	83	

Dividend	7							8							9						
Divisor	1	2	3	4	5	6	7	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	9
No. of occurr.	23	72	55	38	46	32	54	17	318	30	250	22	28	39	91	19	50	124	49	25	15	18	30	38

Dividend	10								11								12				
Divisor	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9
No. of occurrences	258	38	46	120	19	9	24	24	32	21	16	3	7	11	14	3	198	123	152	29	93	9	16	7

Dividend	13								14								15				
Divisor	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9
No. of occurrences.....	45	16	15	11	7	4	5	3	77	20	13	5	8	44	8	6	69	98	16	79	8	8	4	6

Dividend	16								17								18				
Divisor	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9
No. of occurrences...	180	19	130	14	6	9	98	3	61	9	15	14	6	6	12	3	69	49	13	6	28	7	7	23

Dividend	19								20								21				
Divisor	2	3	4	5	6	7	8	9	3	4	5	6	7	8	9		3	4	5	6	7	8	9	
No. of occurrences.....	21	6	10	5	3	4	10	4	24	86	65	11	3	23	5		54	12	8	5	43	10	5	

Dividend	22								23								24				
Divisor	3	4	5	6	7	8	9		3	4	5	6	7	8	9		3	4	5	6	7	8	9	
No. of occurrences.....	17	16	15	8	13	6	15		7	8	11	8	6	3	2		91	76	18	50	5	61	1	

Dividend	25								26								27				
Divisor	3	4	5	6	7	8	9		3	4	5	6	7	8	9		3	4	5	6	7	8	9	
No. of occurrences.....	11	13	105	5	6	5	3		5	6	3	3	4	6	3		46	8	10	4	2	6	25	

Dividend	28								29								30				
Divisor	3	4	5	6	7	8	9		3	4	5	6	7	8	9		4	5	6	7	8	9		
No. of occurrences.....	4	36	8	3	19	3	7		6	8	0	5	11	2	3		21	27	25	6	7	13		

Dividend	31								32								33				
Divisor	4	5	6	7	8	9			4	5	6	7	8	9			4	5	6	7	8	9		
No. of occurrences.....	4	3	1	1	4	2			50	11	3	6	39	5			8	7	7	2	6	1		

Dividend	34																				
Divisor	4	5	6	7	8	9															
No. of occurrences.....	8	3	5	2	1	1															

TABLE VII—(Continued.)

Dividend	35						36						37						38					
Divisor	4	5	6	7	8	9	4	5	6	7	8	9	4	5	6	7	8	9	4	5	6	7	8	9
No. of occurrences	10	31	5	24	5	3	37	16	22	2	6	19	12	8	7	5	3	9	7	8	7	1	1	5
Dividend	39						40						41						42					
Divisor	4	5	6	7	8	9	5	6	7	8	9	5	6	7	8	9	5	6	7	8	9
No. of occurrences	4	3	7	4	3	1	38	9	2	34	2	6	6	3	7	5	7	28	30	10	3
Dividend	43						44						45						46					
Divisor	5	6	7	8	9	5	6	7	8	9	5	6	7	8	9	5	6	7	8	9
No. of occurrences	7	5	10	3	3	7	6	4	5	0	24	6	7	10	20	3	3	2	2	2
Dividend	47						48						49						50						51					
Divisor	5	6	7	8	9	5	6	7	8	9	5	6	7	8	9	6	7	8	9	6	7	8	9
No. of occurrences	6	2	2	0	3	7	17	4	33	2	4	7	27	9	2	4	6	3	8	2	3	1	2
Dividend	52						53						54						55						56					
Divisor	6	7	8	9	6	7	8	9	6	7	8	9	6	7	8	9	6	7	8	9
No. of occurrences	5	5	5	3	4	3	2	2	12	5	1	16	5	3	4	2	0	13	16	8
Dividend	57						58						59						60						61					
Divisor	6	7	8	9	6	7	8	9	6	7	8	9	7	8	9	7	8	9
No. of occurrences	0	3	1	3	2	2	3	1	2	3	0	3	3	9	1	1	2	5
Dividend	62						63						64						65						66						67					
Divisor	7	8	9	7	8	9	7	8	9	7	8	9	7	8	9	7	8	9
No. of occurrences	4	6	1	17	5	9	5	22	0	1	10	1	2	1	4	0	1	1
Dividend	68				69				70			71			72			73			74			75		
Divisor	7	8	9	7	8	9	8	9	8	9	8	9	8	9	8	9	8	9
No. of occurrences	1	3	2	0	6	1	6	2	1	0	16	10	7	5	3	3	5	3
Dividend	76			77			78			79			80		81		82		83		84		85	
Divisor	8	9	8	9	8	9	8	9	9	9	9	9	9	9
No. of occurrences	3	2	3	0	4	1	0	2	4	15	2	4	1	2
Dividend	86		87		88		89																	
Divisor	9		9		9		9																	
No. of occurrences	0		3		2		7																	

some extent by using up time that could be spent better for other purposes, and (though not necessarily) by detracting from the pupil's interest in arithmetic. In certain cases, however, such excess practice and over-learning are actually desirable. Three cases are of special importance.

The first is the case of a bond operating under a changed mental set or adjustment. A pupil may know 7×8 adequately as a thing by itself, but need more practice to operate it in $\begin{array}{r} 285 \\ 7 \\ \hline \end{array}$ where he has to remember that 3 is to be added to the 56 when he obtains it, and that only the 9 is to be written down, the 5 to be held in mind for later use. The practice required to operate the bond efficiently in this new set is desirable, even though it is excess from a narrower point of view, and may make the straightforward "seven eights are fifty-six" over-learned. So also a pupil's work with 24, 34, 44, etc., + 9 may react to give what would be excess practice from the point of view of $4 + 9$ alone; his work in estimating approximate quotient figures in long division may give excess practice on the division tables. There are many such cases. Even adding the 5 and 7 in $\frac{5}{12} + \frac{7}{12}$ is not quite the same task as adding 5 and 7, undisturbed by the fact that they are twelfths. We know far too little about the amount of practice needed to adapt arithmetical bonds to efficient operation in these more complicated conditions to estimate even approximately the allowances to be made. But some allowance, and often a rather large allowance, must be made.

The second is the case where the computation in general should be made very easy and sure for the pupil except for some one new element that is being learned. For example, in teaching the meaning and uses of "Averages" and of uneven division, we may deliberately use 2, 3 and 4 as divisors rather than 7 and 9, so as to let the pupil's energy all be spent in learning the new facts, and so that the fraction in the quotient may be something easily understood, real, and significant. In teaching the addition of mixed numbers, we may use, in the early steps, cases like

11½		79½
13½		98½
24	rather than cases like	67

so as to save attention for the new process itself. In cancellation, we may give excess practice to divisions by 2, 3, 4, and 5 in order to make the transfer to the new habit of considering two numbers together from the point of view of their divisibility by some number. In introducing trade discount, we may give excess practice on "5% of" and "10% of" deliberately, so that the meaning of discount may not be obscured by difficulties in the computation itself. Excess practice on, and overlearning of, certain bonds in thus very often justifiable.

The third case concerns bonds whose importance for practical uses in life or as notable facilitators of other bonds is so great that they may profitably be brought to a greater strength than 199 right out of 200 in 2 sec. or less, or be brought to that degree of strength very early. Examples of bonds of such special practical use are the subtractions from 10, $\frac{1}{2} + \frac{1}{2}$, $\frac{1}{2} + \frac{1}{4}$, $\frac{1}{2}$ of 60, $\frac{1}{4}$ of 60, and the fractional parts of 12 and of $1.00. Examples of notable facilitating bonds are "ten 10's = 100," "ten 100's = 1000," additions like $2 + 2$, $3 + 3$, and $4 + 4$, and all the multiplication tables to 9×9.

In consideration of these three modifying cases or principles, a volume could well be written concerning just how much practice to give to each bond, in each of the types of complex situations where it has to operate. There is evidently need for much experimentation to expose the facts, and for much sagacity and inventiveness in making sure of effective learning without wasteful overlearning.

The facts of primary importance are:

(1) The text-book or other instrument of instruction which is a teacher's general guide may give far too little practice on certain bonds.

(2) It may divide the practice given in ways that are apparently unjustifiable.

(3) The teacher needs therefore to know how much practice it does give, where to supplement it, and what to omit.

(4) The omissions, on grounds of apparent excess practice, should be made only after careful consideration of the third principle described above.

(5) The amount of practice should always be considered in the light of its interest and appeal to the pupil's tendency to work with full power and zeal. Mere repetition of bonds when the learner does not care whether he is improving is rarely justifiable on any grounds.

(6) Practice that is actually in excess is not a very grave defect if it is enjoyed and improves the pupil's attitude toward arithmetic. Not much time is lost; a hundred practices for each of a thousand bonds after mastery to 199 in 200 at 2 seconds will use up less than 60 hours, or 15 hours per year in grades III to VI.

(7) By the proper division of practice amongst bonds, the arrangement of learning so that each bond helps the others, the adroit shifting of practice of a bond to each new type of situation requiring it to operate under changed conditions, and the elimination of excess practice where nothing substantial is gained, notable improvements over the past hit-and-miss customs may be expected.

(8) Unless the material for practice is adequate, well-balanced and sufficiently motivated, the teacher must keep close account of the learning of pupils. Otherwise disastrous under-learning of many bonds is almost sure to occur and retard the pupil's development.

INTELLIGENCE AND ITS MEASUREMENT: A SYMPOSIUM*

VIII. By V. A. C. HENMON,
University of Wisconsin.

1. The nature and measurement of intelligence. Intelligence, in the ordinary acceptance of the term, has been defined by Lester F. Ward to be "intellect coupled with the product of its operation," or in other words, "intelligence is intellect *plus* knowledge." This certainly corresponds to the use of the term in pedagogical and sociological theory and in practical life. The intelligent man is the well informed man and one who is capable of readily appropriating information or knowledge. Without doing violence to language the psychologist can not give it any other technical significance. Intelligence, then, involves two factors,—the capacity for knowledge and knowledge possessed. The untutored savage or barbarian may have high intellectual capacity, but without knowledge we should not ordinarily call him an intelligent man. We could best only say that he had a fine mind, a fine intellect, high intellectual power, or else that he possessed a high native intelligence. But native intelligence from the point of view of this definition is a misnomer, as is also the customary distinction between intelligence tests and achievement tests. Nothing but confusion results from using the term intelligence to designate the level reached in the acquisition of knowledge or mental age, in which sense it is something which can be developed and trained, and using the same term to designate intellectual capacity or mental alertness which is probably given once for all with one's constitution and can not be materially changed by anything we do. Achievement tests while measures of knowledge possessed are in turn indices of capacity for knowledge, though valueless for comparative purposes of crude, uninstructed intellect, only because there has been no guarantee of equality of opportunity. Intelligence tests, on the other hand, as at present constituted, are a mixture in varying proportions of tests of knowledge, for example, the constantly recurring tests of arithmetical reasoning and of information, with those whose novelty is such that specific past experiences function much less readily and there is an approximation toward relative equality of opportunity. In any case, both groups of tests are measures of both native and acquired intel-

*Continued from the March, 1921, issue. In that issue appear statements by Doctors Thorndike, Terman, Freeman, Pintner, Colvin, Ruml and Pressey.

ligence with differences in emphasis. The high correlations between intelligence tests and achievement tests and the high correlations between intelligence tests and teachers' estimates or scholastic standings indicate their essential similarity. Teachers' estimates and school marks, which are used as measures of reliability and validity of intelligence scales, in default of any absolute standard of reference, are based on what children know and on their capacity to know.

Intelligence ought to retain this broad significance. To define it with Stern as "a general capacity of an individual consciously to adjust his thinking to new requirements," which is substantially James' definition of reasoning, or with Meumann as a general capacity for "independence, originality and productiveness in thinking" is to narrow its meaning arbitrarily and to neglect its compound nature. The scholarly or erudite man who has merely acquired the knowledge created by others may not represent as high a degree of intelligence as one who is independent, original and productive in his thinking, but we should scarcely say that he is unintelligent. Intelligence is indicated by the capacity to appropriate truth and fact as well as by the capacity to discover them.

What we want of course, is a measure of crude intellect or mental alertness as such, a scale for intellect not a scale for intelligence. Any test which requires mental processes, be they either on the sensory-motor level or on higher levels, which measures the sensitiveness, responsiveness, and retentiveness of the nervous system, is a test of intellectual capacity. It is only because of the absence of an absolute standard of reference, the inability to discount the effect of past experiences, and the narrow conception of the meaning of intelligence, that the tests of higher mental processes appear to be better measures of intellect, and not because they involve intellectual capacity any more evidently. The essential thing in a test of pure intellect is to secure such novelty as to reduce inequality in opportunity and training to a minimum, and such complexity as to reveal differences clearly and unambiguously. It is not a matter of the kind of mental process involved primarily. A properly constituted test of intellect ought, therefore, to take a wide sampling of so-called lower as well as higher processes and the relative weights to be assigned to each element determined. The tests must be extended to measure the capacity to learn in other directions than the ability

to use words, signs, and symbols. The so-called general intelligence tests are not general intelligence tests at all but tests of the special intelligence upon which the school puts a premium. The extension of the term intelligence to include what Thorndike has called mechanical intelligence, the ability to manipulate things, and social intelligence, the ability to manage men, means that we need not one, but several scales of intellect from which we can secure an intellect profile. We have been gradually evolving a type of test for abstract intellect that is of use in classification and prognosis in school, but is of no convincing value as a measure of general intelligence as such, nor of demonstrated value in vocational guidance and direction.

2. *Next steps in research.* The construction of group tests proceeds apace. Many of them are uncritically and hastily assembled, hurried into print without norms or standards to satisfy the demand of some publishing company, and often merely rearrangements of familiar material, selected on the basis of the author's opinion of the merits of the component tests. While experimentation in test construction is desirable, critical evaluation of the more important group tests is far in arrears and much needed. General mental age determinations from existing tests are seriously open to question; the intelligence quotient derivatives involve questionable assumptions as to the nature of mental growth; the relatively narrow range of functions measured in many teams of tests make them invalid as measures of general ability; the assignment of equal weight to each test and to each item in the individual tests is surely wrong. These matters need careful study with the better known and better standardized tests.

The correlations between the scores in many of the group scales as a whole and such standards of reference as we possess do not differ very materially. The individual tests, however, differ among themselves considerably in diagnostic and prognostic value, in susceptibility to practice, in interest to the one tested, in ease of administration and scoring, in relative emphasis on previous training and capacity for meeting relatively new situations. What is needed clearly is more accurate scaling of the items within the individual tests, the reduction in number or elimination of tests that test apparently the same or similar functions, the application of the method of partial correlations to determine causal relationships and the correct weights to be assigned to the individual tests in the

scale as a whole. Besides the refinement in the tests of abstract intellect themselves we need also to determine the importance of various character traits which apart from intellect as such make for success in the tests. Until these things are done we shall be in constant danger of misusing the results of tests for classification and prognosis even in schools. We are still far from being able to give the significance and meaning of a test score.

Besides the need for critical evaluation, standardization and interpretation of tests of abstract intellect stands the need for tests of ability to manipulate things and the ability to deal with men, leadership and adaptability in social relationships to complete the intellect profile suggested above.

IX. By Joseph Peterson,
George Peabody College for Teachers.

1. The nature and measurement of intelligence. Intelligence seems to be a biological mechanism by which the effects of a complexity of stimuli are brought together and given a somewhat unified effect in behavior. It is a mechanism for adjustment and control, and is operated by internal as well as by external stimuli. The degree of a person's intelligence increases with his range of receptivity to stimuli and the consistency of his organization of responses to them. In certain higher learning processes the bright subject more quickly and more systematically enlarges the range of his cues or 'guide-posts,' and consequently more readily masters the situation than does the dull one, who must come back again and again to the more immediate elements. The latter cannot keep in mind the complexity of the situation. This difference is found to hold for children of very different years but equal mental ages, contrary to the results of a well-controlled experiment by Woodrow, one, however, the results of which seem to me to have been generalized too much by him and by some other writers. A dull and a bright child of equal mental age have gone through equal mental-age levels, but the bright child seems to have made the same acquirements in the lesser time.

One's emotionality, responsiveness to abstract symbols, degree of energy and of perseverance, etc., certainly condition one's intelligence. We cannot analyze such differences here, but can only suggest that equal mental ages by any test do not indicate just equal

intelligence for all situations. We are probably too prone in our practical work of testing to assume equality of intelligence in such cases, since our attention is taken up with the one kind of reaction most susceptible to testing. General intelligence, if it is a reality at all, is probably not a separate or constant factor, but a composite of many different abilities, and probably means different things in unlike situations, as different abilities are stressed. Such factors as energy and perseverance, degree of disturbance by emotions and self-consciousness, and many others that play their rôles in one's success in life, have not yet been successfully brought into the field of measurement by tests, especially by group tests.

We have moved a long way from testing intelligence by specific sensory discriminations, reaction time, etc., and have rightly come to emphasize constructive imagination, quickness of the perception of relations, and other higher mental functions; but so far as the production of group intelligence tests that will measure abilities making for success in many phases of practical life is concerned, we have not yet arrived at an entirely satisfactory solution. Undoubtedly we shall need different tests for different purposes, though general intelligence tests seem to be the most promising on the whole; but our chief problem is to present the suitable problematic situation. Reactions with a pencil to highly abstract symbols certainly constitute an inappropriate type of test for an adult, let us say, who has spent years developing specific habits of other kinds. Reactions to pictures are far from being suitable substitutes for the more concrete situations demanded in the case of illiterates, although with proper precautions picture tests can be made very serviceable. Other objective materials are not yet very suitable for group tests.

2. *The next step in group mental testing.* The "next step" will probably not be a radical departure from present methods. Some refinements in method that seem to me probable and desirable are:

1. More attention to the equality of score points and to the treatment of errors. At present the score value of different test-units is largely a matter of personal judgment. Some 'scales' increase in difficulty more or less, uniformly in successive units, but usually even such increases are not accurately evaluated. In the most widely used scales units of increasing difficulty are found in some of the tests, while in other tests the units are of approximately equal difficulty. Even though fairly satisfactory results are thus

obtained, especially in the selection of individuals of extreme abilities, it would be sanguine to hope for much development in the science by such methods. No one knows just how much we fail with individuals of more nearly median ability, who differ from one another by smaller amounts; and to find this out requires more perfection in the criteria by which we usually evaluate tests. The development of methods for obtaining more reliable criteria is much needed, otherwise one test is correlated with another and the other one is in turn evaluated by it.

In the scoring of errors Thurstone has made a good start, but his method presupposes linear regressions and approximate equality of score units. An illustration may show its value. If we use the total score in the Pressey Cross-Out Test, Schedule E, as criterion—the only one available—and seek by Thurstone's formula to find the best scoring value of C in the equation $S = R + CW$, in which S stands for score, R for rights and W for wrongs, we get from tests of 53 12-year-old unselected children in Nashville schools, $C = -.44$, $.69$,-24, and .21, for tests 1, 2, 3, and 4, respectively. Using these values of C, we get the following correlations of these four tests, in order, with the criterion: $rI(R + CW) = .61$, .96, 53, and .63. The corresponding values of r between total scores and rights (neglecting errors as was done in the derivation of the standards) are .55, .71, .48, and .69. This example shows that different parts of the same test may have, and probably should have, different score methods. These the authors of future tests should determine, and they should prepare tables to facilitate rapid scoring. Such treatment of group tests, following careful evaluation of score units, would greatly increase their usefulness, but it requires reasonably reliable criteria.

2. More thorough-going analysis of correlation coefficients. In the first place, we must not forget that $r = \sqrt{b_1 b_2}$, and that therefore it may be very deceiving in itself. There are also many factors making for non-linearity of regression lines that need careful study, many of them being due to such imperfections in the test itself as improper time allowance, improper gradation as to difficulty, and poor scoring methods. Formulae for non-linear relationships seem to me to have very little practical value for interpretation. We have found it more profitable to make distribution tables of scores and to study the actual curves through the means of rows and columns than to use these formulae for anything but merely to show the de-

gree of relationship. Such an analysis may, for instance, show that a test has significance only for certain degrees of mental ability, as at extremes, a condition that the correlation ratio alone could not reveal.

It should be noted also that correlations have very little significance for comparisons with others unless the degree of heterogeneity of the group tested is indicated. A low correlation may in some cases be better than a high one in others. There is little significance in the value of r itself. Partial and multiple correlation methods will doubtless be used in increasing measure both in the selection of suitable tests and in the analysis of results. Another illustration from the test of twelve-year-old children may show this better than will any other method. The four tests—counting rights only—showed inter-correlations as follows: $r12 = .33$, $r13 = .31$, $r14 = .20$, $r23 = .26$, $r34 = .54$. The partial coefficients of the second order are $r12.23 = .22$, $r13.24 = .22$, $r14.23 = .02$, $r24.13 = .49$, $r34.12 = .03$. It is seen that tests 1 and 4 have nothing in common not included in 2 and 3, and that 3 and 4 also test different functions but for the effects of 1 and 2. On the other hand, 2 and 4 have much in common not included in the others. (This analysis is based on Pressey's method of scoring, however.) On the basis of facts such as these it is possible to select a battery of tests that measure widely different functions, and thus to get better general intelligence tests as well as valuable data on the interrelation of mental functions. A better scoring method in this case might reveal different relations from those here indicated, but we are interested only in the method of procedure.

In the April number of this journal I have indicated some important problems relative to the time element in tests.

X. By L. L. Thurstone,
Carnegie Institute of Technology.

What is meant by intelligence?

Intelligence as judged in every-day life contains at least three psychologically differentiable components: a) the capacity to inhibit an instinctive adjustment, b) the capacity to redefine the inhibited instinctive adjustment in the light of imaginally experienced trial

and error, c) the volitional capacity to realize the modified instinctive adjustment into overt behavior to the advantage of the individual as a social animal.

The inhibition of an instinctive adjustment involves the substitution of an overtly passive deliberative attitude against environmental, social, and instinctive pressure. The degree of intelligence is partly measurable by the relatively early inhibition of the instinctive adjustment at a stage when it is not yet defined.[1] Conceptual thinking is rendered possible by the inhibition of the instinctive adjustment at a very early stage of its formation. A concept is an unfinished act inhibited before it has become personally concrete and it therefore has more extension and less intention than the idea which is an act inhibited after it has become motorially more defined. The capacity to redefine the inhibited instinctive adjustment in the light of imaginally experienced trial and error is essentially the analytical element of intelligent work. In routinized intelligent work the mental trial and error experience may be rationally controlled, but in unroutinized intelligent work the mental trial and error experience has its causal antecedents in loose affective and conative analogies. These subconscious analogies of intelligent work constitute alternatives of less resistance at a preconscious stage of the unfinished, undefined and inhibited instinctive adjustment. Analytical capacity may be thought of as flexibility of affective and conative associations which precede the cognitive phase in the definition of the stimulus into the response.

These three phases of intelligent behavior may be labelled for convenience a) inhibitive capacity; b) analytical capacity, and c) perseverance. One may possess the inhibitive capacity without keen powers of analysis. This leads occasionally to an appearance of intelligence by silence. Such a mind is often capable of recognizing and adopting the successful analyses of other minds without being able itself to produce the analyses.[2] The inhibitory capacity for the early, abstract, impersonal stage of instinctive adjustment does not necessarily coexist with inhibitory capacity for the later more defined and personally significant stage of the instinctive adjustment.

---

[1]By "early" I do not refer to childhood or the education of youth. I mean by "early" the incipient response as it is being consciously particularized in every moment of normal waking life.

[2]When such a mind is by superior social traits adapted for administrative work it will lean on competent assistants for analyses. This type of mind is common in industry and in a deplorably exaggerated form in politics. It was described even by Machiavelli in "The Prince."

Hence a man may possess inhibitory capacity for abstract impersonal thinking and lack self-control for those instinctive adjustments which have reached uninhibited a personally and socially significant and overtly defined stage. A man may possess considerable inhibitory and antlytical capacity and still be socially quite ineffective because of insufficient volitional drive to realize his intelligence into consistent action. Such a mind is not infrequently the subject of ridicule from those who would justify their own mental worth in spite of modest analytical powers.

Analytical capacity concerns the preconscious and early conscious stage in the definition of instinctive adjustment. It implies the capacity to inhibit instinctive pressure at that stage so as to allow flexibility of definition of the response. This has been picturesquely called by Knowlson free interplay with the subconscious.[a] It is only natural to expect that such inhibitory capacity should appear when the instinctive drives are relatively weak. Hence superior analytical ability is not infrequently associated with inferior instinctive energy in its nutritional, sexual, gregarious, and self-assertative forms. When such an individual discovers his inferior instinct equipment he may center his self-satisfaction on his analytical powers as a compensation. Such individuals need the assistance of other volitionally stronger minds in order to make their labors socially effective.

When the analytical capacity is associated in the same mind with a fair degree of instinctive and volitional energy we have the bright person in practical life. There is a fundamental distinction between bright and profound intelligence. The mind which possesses a fair amount of inhibitive, analytical, and volitional capacity will in general aim for effectiveness in the immediate present. Such a mind is bright. When, however, a mind is more generously endowed in all three of these respects or when it is unbalanced by reduced volitional capacity, it deals with the earlier and more abstract relations in its adjustments. Such a mind is profound and its analyses concern temporarily remote ends. I am quite sure that our intelligence tests in which the candidate races against time for a few minutes do not measure adequately the more inhibited and deliberative profound type of intelligence. It does fair justice to the more frequent well balanced bright mind.

---

[a] "Originality," by T. S. Knowlson.

I should summarize my definition of intelligence as follows: Intelligence is the capacity to inhibit instinctive behavior in an unfinished stage of its formation and to modify it at that stage by means of an imaginal stimulus which is relatively remote from that which is immediately and perceptually present. The imaginal stimulus in intelligent behavior is of preconscious origin. It becomes conscious as an alternative to be controlled, accepted or rejected rationally, but its origin is a preconscious, irrational, uncontrolled association of affective and conative similars. This definition is consistent with that of Baldwin in his discussion of meaning in which he says: "It is in the passage from the bare recognition of each item presented as being just what it is, to its treatment as being in some sense not what it is, but what it may become or be used as, that psychic meanings as such arise." Bergson says that "intelligence, considered in what seems to be its original feature, is the faculty of manufacturing artificial objects, especially tools to make tools, and of indefinitely varying the manufacture."

There are two features in intelligent behavior that I have tried to emphasize, namely, that it implies inhibition and that it consists largely in rendering conscious an unfinished act. The sooner the act becomes conscious, the more crude and unfinished it is when it becomes conscious, the greater will be the range of possible overt behavior into which it may become particularized, and the higher will be the intelligence of that particular moment. I have previously described this point of view with reference to the momentary psychosis.[4] The element of inhibition in intelligent behavior may be partly justified if thought of in relation to its various ramifications such as the inhibitory functions of the cortex, the recession of the stimulus in intelligent behavior,[5] the frequently drawn contrast between intelligent behavior and instinct- or habit-determined behavior, the modification of instinctive protopathic behavior into intelligent epicritic behavior,[6] and the analytical intelligence of the inhibited introvert as contrasted with the analytical impatience of the extrovert.

By what means can intelligence be best measured by group tests? This question can be given at least three different interpretations.

---

[4] "The Anticipatory Aspect of Consciousness," Journal of Philosophy, Psychology and Scientific Method, Oct. 9, 1919.
[5] Holt: "The Freudian Wish and Its Place in Ethics."
[6] Rivers: "Instinct and the Unconscious."

We may set out to measure pure intelligence as a trait which is only partly responsible for social success. In this case we shall probably have no criterion by which to judge the test. We may set out to measure brightness using as our criterion the estimates of avowedly competent judges of brightness and dullness with the necessary admixture of volitional and emotional ingredients. We may set out to measure ability to do a specified practical task for which we assume that intelligence is an indirect partial requisite. This latter type of mental measurement is the customary one. Our criterion is the relative success of the candidate in school, college, salesmanship, clerical work, or what not. Any test procedure which has diagnostic value with reference to a specified criterion is acceptable. It is much more difficult to measure intelligence as such and I am not sure that it has ever been done. For administrative purposes, we do not need to measure pure intelligence. In fact our diagnoses are probably more effective if we test other significant factors in our tests, although it would be scientifically interesting to be able to tease out the intellectual, volitional, affective, social, heredity, and experiential elements.

It would be well if we could draw a sharp line of distinction between service and research in mental test work. These two purposes are often confused. The results of research may be summarized for direct application in some mental test service. Some research data may be obtained from service records but the reverse is rarely successful. Some mental testers are justifying their use of a test in test service simply because they have norms for it. This should never be done. No test should ever be used for any kind of service unless it is known what it is that the test diagnoses. This necessitates a criterion for every test. If the test correlates well with a criterion such as chronological age, it can be used to determine age. But a test may be good for one criterion and poor for another criterion. We should never talk about a "good" test without telling what it is that the test is good for—namely, the criterion. Another relatively common mental test illusion is that if a new test has a high correlation with the Army Alpha, and if the Army Alpha is good for some specified criterion, the new test is forthwith assumed to be good for that criterion too. This does not follow. The only safe way is to take the trouble to determine the diagnostic value

of the new test without inferring it through another test as an intermediary.'

It has always seemed to me rather peculiar that we measure children's intelligence in terms of their chronological age. I should like to suggest that possibly the underlying assumptions would be less troublesome if we stated the intelligence of an eight-year old in terms of his percentile rank with reference to other eight-year olds, or in terms of the standard deviation of test scores for eight-year-old children. From such a direct rating one could of course readily ascertain that chronological age the mean intelligence test score of which any given child attains.

What are the most crucial next steps in mental test research? Now that we have accumulated enough norms and correlation coefficients to make one dizzy it is about time that we begin to formulate some mental test principles. Our literature abounds with instructions for giving particular tests and with statistical data of administrative interest. But we have so far no science, no principles, no psychology in the mental test literature. We have done well with our empirical cut-and-dry methods in mental test work. We can do better if we rationalize our findings in this field. Unfortunately there are relatively few mental testers who are interested in deriving psychological fundamentals from mental tests.

I should like to see another line of mental test work opened up, namely, the diagnosis of the volitional and emotional characteristics which determine our character traits. Intelligence is only one of the elements in mentality and it has been overworked because of being accessible to measurement. We should investigate the possibility of diagnosing character traits by some new kinds of mental test, self-analysis forms, and other procedures. Who knows but that individual differences in the various characteristics of our reflexes may be diagnostic of character traits? Why is it that executives are in general taller and heavier than other people with similar habits of life? Why is it that we guess right more often than wrong about the character traits of a stranger by merely looking at his face? These questions should be rationalized as science. What are the possible schemes by which we may successfully classify character traits? What types of mentality are produced by the combination of intelligence and different emotional and volitional traits?

'Yule: "Introduction to the Theory of Statistics," p. 250.

By what physiological and psychological technique can we diagnose these types? What are the mental processes that distinguish special ability? Can these be taught to others? These are worthy problems for mental test psychologists.

It is high time that we quit justifying ourselves as psychologists by simply standardizing mental tests. If we attack the individual diagnosis of character traits as energetically as we have been giving group tests the results will be of far reaching psychological, educational and social significance.

XI. By Herbert Woodrow,
University of Minnesota.

1. The nature and measurement of intelligence. In attempting to state the meaning of intelligence, I should say first that it is an acquiring-capacity. One's actual, present capacity for doing any particular thing is obviously dependent upon his training and learning, or, in general, upon his past experience; whereas, in the measurement of intelligence, the attempt is always made to minimize the effects of past experience by utilizing performances, the acquisition of the capacity for which is presumably favored equally by the past environments of all the individuals measured. When, for example, we use a child's present capacity to count backwards from twenty as a partial index of his intelligence, the assumption is, clearly, that the past environments of all children have been equally favorable for the acquisition of this ability. If this ability was one which could be readily taught, and, further, one which was taught some children but not taught others, it would be valueless as an index of intelligence. The present capacity to count is a measure of intelligence only in so far as it is an index of the capacity to *learn* to count, that is, the capacity to acquire the capacity to count. Intelligence, then, is the capacity to acquire capacity.

Some psychologists, no doubt, would urge the desirability of defining intelligence as a capacity which is fixed at an early age and is largely innate. These, while conceding that intelligence is a thing that *grows,* might argue that the capacity to acquire is, itself, largely acquired. They might point out that, at the age of fifteen,

for example, one has the capacity to acquire abilities that could not possibly be acquired at the age of six, and that the difference in the capacity to acquire at these two ages was largely due to what had been learned in the years between them. I think this argument is sufficiently answered, however, by regarding the acquisitions between any two ages as progressive steps in the acquiring of an ability, that is, as part of the learning of whatever is learned easier as a result of them. I should, therefore, incline to the view that it is not a necessary part of the *definition* of intelligence, but simply an established fact, that intelligence, except for being a *growing* thing, is fixed partly by heredity and partly by environmental factors acting before the age of five. I hold, further, the view which I proposed in this Journal some years ago (1917) as an experimental conclusion, that the growth of intelligence is neither produced, nor appreciably accelerated, by learning.

Now, if intelligence is to be defined as a capacity to acquire capacity, it is imperative to define the acquired capacity. I believe this capacity may be correctly described (using neither purely mental nor purely behavioristic terms) as the capacity for such mental activity as is most effective in bringing about success, or, in slightly different phraseology, as the capacity for success in so far as success is dependent upon mental processes, either past or present. By success, is here meant, not success in any one line of endeavor, but a theoretical success, namely, success in all those innumerable performances which may be regarded as desirable. The capacity for such a theoretical success would have as its practical result success in coping with any situation in which the individual might find himself.

This conception of intelligence implies that mind is an instrument for the securing of desirable ends. It implies that between mind and body there exists a relationship—a fact which I believe no living psychologist denies—but does specify the ultimate nature of the relationship. Intelligence cannot be defined in purely mental terms, because the capacity for acquiring valuable modes of mental functioning is itself not mental; it is certainly, in part, a matter of the condition of the brain. Moreover, success can be attained only as the immediate result of behavior. All "mental measurements," even those concerned with "intensity of sensation," are primarily measurements or descriptions of behavior under measured or described environmental conditions. On the other hand, intelligence cannot

be satisfactorily described in purely behavioristic terms, because (I mention only one reason, one which is persistently overlooked by behaviorists) while the degree to which behavior is intelligent is simply the degree to which it attains success, success has no real meaning except by reference to some want, desire, intention, plan, or purpose.

From the above point of view, the problem of the relative desirability of different kinds of testing materials is essentially the problem of the relative value of the different kinds of mental processes in securing "success in general." The existing data show that many of the "simpler" mental processes are of value, as well as all, or nearly all, of the more complex ones. In general, however, the more complex processes are much the more valuable. Fine pitch discrimination, for example, while by no means useless, is of far less value than good attention and correct reasoning. From these considerations, it follows that any accurate measurement of intelligence, whether by individual or group tests, must call into play both the more complex mental processes and the simpler ones, but should invoke each, or weight the measurement of each, in exact proportion to its usefulness in securing "success in general." Since, as a rule, the more complex processes are the more useful, they should, as a rule, be given the greater weight.

Now the relative usefulness of the different forms of mental activity varies with sex, social status, stage of civilization and, most important of all, with degree of intelligence and therefore with *age*. The best tests, the tests which go farthest in determining the outcome of the measurement of intelligence, will differ with the age of the children measured, and, in general, will be tests of the simpler mental functions in young or unintelligent children and tests of the more complex functions in older or more intelligent ones. To use mainly tests of simple mental functions at the early ages, and mainly tests of more complex ones at the later ages, is simply one way, though probably the most convenient one, of giving greater weight to the measurement of simple mental functions in the early stages of the scale. However, even a scale designed solely for children of some specified early age will involve the measurement of capacities which, though perhaps all comparatively simple, will vary to some degree in complexity and in value; and even such a scale should, therefore, in accordance with the general rule stated above, give weight to the

measurements of different capacities in exact proportion to their usefulness.

2. Next steps in research. While I am unable to outline any "next steps" which I should call "crucial," there are numerous desirable ones. I should urge, in the first place, the desirability of repeating again and again the steps already taken. Investigators who have worked extensively in this field have, after all, been few; and the very fact that they have done so well entitles them to successors who can devote years of patient toil and the necessary originality to the devising of new methods of measurement; to a careful analysis of the mental processes involved in each test, together with the study of individual variations in the same; to the study of the relative weight of different tests as measurements of intelligence; to establishing norms, not from data furnished by untrained people, distorted by the "personal opinion" and "benefit of the doubt" methods of scoring, but from data obtained in a scientifically reliable manner; and, in short, to devising tests with a better logic behind them, both from the viewpoint of psychological theory and the mathematical theory of measurement.

More information is needed on every point connected with brightness, or relative intelligence, in distinction from absolute intelligence. No doubt further data will soon be forthcoming on individual variations in brightness with chronological age, with anatomical age, with disease, with harmones, with learning and with various environmental factors. In the measurement of brightness, even more than in the measurement of intelligence, we need further work in the logic of method. For studying the relation of brightness to anatomical age, we need, and we shall have, a much more adequate scale for the measurement of anatomical age than any now in existence.

XII. By W. F. Dearborn,

Harvard University.

1. The commonly accepted definition of intelligence such as the capacity to learn or to profit by experience evidently involves a description of the nature of the individual and his reactions to the environment, which, I believe, can now be adequately given in terms of the current "objective" psychology. The older conceptions of per-

ception, association, imitation, and reasoning are described concretely in terms of "unconditioned" and "conditioned" response, and learning, as the development of intelligence, appears as a continuous process from the establishment of the first conditioned response to the "discriminating perception" called reasoning.

2. Theoretically, it would follow that measurement of the actual progress of representative learning would furnish the best test of intelligence. For practical reasons most tests now in common use are not tests of the capacity to learn, but are tests of what has been learned. The assumption is made that if one samples the results of learning in matters where all the individuals tested have had an equal chance at learning, he may arrive at an estimate of the capacity to learn. But, since it is difficult to find even simple experiences which are common to all individuals of a given age period, actually, again, one tries by sampling a large range of fairly common experiences to strike an "average" which, despite the fact that a given individual may have missed this or that experience, will still be representative of the individual's learning. The shortcomings in the working out of these assumptions are generally recognized. An extension of this method would still seem desirable and the development of methods of group testing will make possible the more rapid trying out and evaluating of the tests that are adapted to group use. The best of these should be used to extend and make more representative the individual tests now in use. A wider sampling of experiences is clearly necessary in the upper age periods. In addition, individual tests involving actual learning rather than of the results of learning are needed.

3. The methods of determining the mental "age" or status of an individual need revision. The mental age is the resultant of at least three factors, native intelligence, physiological maturity and "environment." An analysis of these factors may be secured by the repetition of mental and physiological tests during a period of years on the same individuals. Assuming for a given group of individuals a constancy in the first and last factors the variations in the second factor, particularly during the adolescent period, will certainly show some variability in the individual intelligence quotients.

4. Finally, the whole question of the unit of measurement assumed in mental age and intelligence quotient needs careful canvassing. Pearson has stated, as a result of his study of the somewhat

limited material of Jaederholm, that a year's mental growth is for a number of years practically constant and equal to the standard deviation of the groups. The conclusion has been drawn from Bobertag's study of the differing amounts of overlapping in the 7th and 8th years as compared with the 11 and 12 year olds that "a year of mental age at the younger ages means a bigger change than it does at the higher ages" (Woodrow). Both statements assume a normal distribution and a fixed variability. The latter assumption is certainly open to question. And, suppose that growth is comparable to the practice curves of learning, where a small gain towards the end of the learning seems harder to make than a larger gain at the start, is this more than saying that your arbitrary units of measurement are changing? Careful use of the results of the group methods of testing should give data for the attacking of these problems.

XIII. By M. E. Haggerty,
University of Minnesota.

1. Intelligence and its measurement. In responding to this request to participate in the symposium on intelligence I take it that what is desired is a more or less dogmatic confession of faith rather than a systematic effort to support by evidence and argument the beliefs expressed.

1. In my thinking the word intelligence does not denote a single mental process capable of exact analytic definition. It is a practical concept of connoting a group of complex mental processes traditionally defined in systematic psychologies as sensation, perception, association, memory, imagination, discrimination, judgment and reasoning. The nearest equivalent to this working concept is James' "differentia of reasoning" as "the ability to deal with novel data" and Dewey's well known "forked-road" theory. This limitation of the concept of intelligence to the "novel-situation" category, so useful for theoretical purposes, however, seems to me too thin and lean for practical purposes of controlling behavior. I prefer to enrich the concept with all the fatness afforded by the inclusion of the more elementary and subsidiary processes instrumental to the crucial act of thinking or reasoning. For the most part, I would

exclude from the concept, emotions, instincts, will-activities and so called character traits.

2. Intelligence is a dynamic concept. It is descriptive of behavior and not of static component parts of the "mind." When one conceives of the mind as a "state" or a "structure," the word intelligence becomes meaningless. The implication of *activity* is essential in the concept of intelligence.

3. Intelligence is native and peculiar to an individual. In this sense it is an instinct defined as an inherited or congenital characteristic. It may be conceived of as akin to, if not dependent on, or a phase of, the native quality of the physical organism, the brain and nerves. Because of the original differences in intelligence or native capacity the same environment affects differently different persons. As a result two persons subjected to the same environment, say the formal procedures of the elementary school, emerge with different mental content, different habits, different interests. These eventuating differences are the criteria by which we may infer the quality of the native endowment of the individual. Often these differences among individuals are less important in themselves than they are for such inferential purposes. Present attainments are important in showing what an individual can now do. In so far as such attainments may be made the basis of inference as to native endowments they are of value for prognosis. Because of their superior prognostic value intelligence tests have a value which mere achievement tests do not have.

4. Intelligence is not merely a qualitative concept. It may be quantitatively considered. When a child has little of it he does poorly in school, and in cases of extreme deficiency he is largely uneducable. When a child has more intelligence he is capable of learning and individuals gifted with great intelligence may learn most rapidly. Similarly in actual life low intelligence means simple occupations and crude civilization. High intelligence means the possibility of efficiency and leadership in the more complex problems of civilized society.

5. Because intelligence is capable of quantitative considerations it is theoretically measurable. Measuring devices with a considerable validity have within recent years been made available in the form of individual and group scales and tests. These scales and tests vary greatly in value, but in increasing numbers they are

meeting essential statistical criteria and through the improvement of the technic of administration and interpretation they are becoming widely usable and useful.

The most satisfactory tests so far devised attempt to measure intelligence in its essential complexity. The earlier attempts to analyze the factors of intelligence into simple elements such as sensation, perception, memory, attention, etc., largely resulted in failure to secure useful tests. Over-analysis apparently distorts the essential phenomena to be measured and to follow this path of experimentation is to court futility. The practically useful test must evaluate the capacities of an individual in the essential complexity necessary to meet actual living situations.

6. The most satisfactory intelligence tests so far devised very largely involve the use of language. Of these the most dependable are tests of a knowledge of vocabulary, of general information, tests for facility in thinking logical relations such as opposites and analogies, for ability in the solution of arithmetical problems, in the completion of elliptical sentences, etc. It is possible and probable that such verbal tests will always be the best and most generally useful means of testing the higher ranges of intelligence. Such verbal tests, however, have their limitations in testing illiterate or partially illiterate adults and children, in testing children in the pre-school years, and in testing such individuals as may find the most adequate expression of their intelligence through manual or other non-verbal activities.

7. An intelligence test is not an end in itself. It is an instrument for diagnosis. It is intended to aid a capable and intelligent person to secure a more accurate determination of the intelligence of a human being than he could secure without such a test, to secure it in a shorter time and in general to enable such a capable and intelligent person to deal with human beings in a more discriminating fashion than he otherwise could do. The score on an intelligence test, however, is not to be regarded as a complete substitute for personal judgment and common sense on the part of the examiner. The analogy most expressive to my way of thinking is to compare the intelligence test in the hands of a psychologist or educator with a thermometer in the hands of a physician. In a similar fashion the test is an instrument for recording symptoms. Its results are one aid in diagnosis.

8. Where education pretends to be based on fundamental principles it must take account of the factor of intelligence. A measure of the intelligence of pupils must, therefore, be available to the teacher, the supervisor, the administrator and possibly to the pupil himself. The extension of intelligence examinations to the public schools under conditions which properly safeguard the interpretation of results seems, therefore, a legitimate movement. The time will probably come when all progressive schools will record the intelligence score of a pupil with the same care that it records his chronological age. For educational purposes, it is more important.

9. Students of experimental education must take account of the intelligence of experimental subjects as they have not done in the past. Most of the studies designed to compare one method of learning with another are deficient in this respect. In many cases the varying intelligence of the subjects is a more determinative factor in the results than is the method of learning used. Indifference to this factor has given us experimental results which are non-interpretable. Most of the experimental work in this field must be done over if it is to throw any real light on educational problems. The freedom with which certain experimentations have sailed along, indifferent to the factor of intelligence, casting ashore their partial findings, has done and is doing much harm.

10. When intelligence tests are used the results must be interpreted in the light of other data indicative of native capacity. In the case of school children such "other data" are school marks, teachers' estimates, school progress, etc.

11. Group intelligence examinations of thirty minutes' duration give a highly useful rating of school pupils. When supported by other data the results can be used as a basis for classification. The value of such a test will be increased by the repetition of the same or a similar test. Individual examinations should be employed in complicated cases.

12. Intelligence is not the only factor conducive to success. Industry, perseverance, loyalty, cheerfulness and other non-intelligence traits are important and in many cases determinative of success.

2. *Next steps in research.* The direction of experimental progress in dealing with intelligence problems seems to me about at follows:

a. The perfection in technique and statistical criteria of verbal tests for the ranges of ability where such verbal tests may be used. The general lines for work are clear; refinement of method is the end sought. Much remains to be done.

b. The development of non-verbal tests for young children, for illiterate and non-English reading children and adults, and for the examination of those special aspects of intelligence,—if any such exist—which are not properly measured by verbal tests. To some degree this means refinement of tests which have already demonstrated their usefulness. It probably also means the qualitative exploration of a new field.

c. The problem of special aptitudes is unsolved. Some evidence exists to show that skill in the analogies tests is prognostic of success in the formal freshman high school algebra. Possibly a number of our so-called general-intelligence tests have such specialized prognostic value. If so, the fact is worth experimental demonstration. It is highly probable that other aptitudes exist which none of our present tests detect. Such an aptitude might be a predisposition to learn telegraphy. It is altogether possible that such special aptitudes are sufficiently evident in childhood and youth to make their study and measurement of important educational significance. Certainly their quest is as inviting as many another which has issued fortunately.

d. Devices for securing the judgment of teachers on the capacities of pupils are undeveloped. "Rating scales" for pupils, for teachers, and for adults give promise of great usefulness if they can be redeemed from their present chaotic state and placed on a scientific basis. By their means we may be able to improve our evaluation of many non-intelligence traits not now subject to objective measurement.

e. The work of Dr. Downey on the "Will-Profile" suggests that the objective measurement of non-intelligence traits is possible. Probably nothing would better supplement our intelligence examinations than would the perfection of an objective measure of the so-called character-traits.

f. The establishment by some scientific organization enjoying public confidence of the essential criteria which satisfactory tests should meet would go far to reducing our discussions to a common language. The current proposal of the National Association of the Directors of Educational Research is a step in the right direction.

THE RELATIONSHIP BETWEEN EYE-PERCEPTION AND VOICE-RESPONSE IN READING

G. T. BUSWELL,
University of Chicago.

PROBLEM.

In oral reading the eye always moves along the line of print in advance of the voice, at times keeping very far in the lead and at other times very little in advance. A mature reader tends to maintain a comparatively wide average span between the eye and the voice, which at times may amount to the space occupied by seven or eight words. An immature reader, however, tends to keep the eye and voice very close together, in many cases not moving the eye from a word until the voice has pronounced it. Reading of this type becomes little more that a series of spoken words because there is no opportunity to anticipate the meaning in large units. An eye-voice span of considerable width is necessary in order that the reader may have an intelligent grasp of the material read, and that he may read it with good expression. If words are encountered which are spelled alike but pronounced differently, such as "read" (present tense) and "read" (past tense), the correct pronunciation and meaning cannot be determined in many cases until the eye has observed the context by looking ahead. The need for a wide eye-voice span is also emphasized when marks of punctuation are encountered. The failure of immature readers to respond to a question mark by a rising inflection of the voice is clear evidence that a narrow eye-voice span has kept them unaware of its presence until it is too late to modify their expression.

In order to determine more fully and accurately the nature of the eye-voice span an investigation was organized to cover a series of problems which were involved. This article presents a brief summary of the experiments, a complete report of which has been made in a separate monograph.[1]

The experiments were grouped into three main divisions. In the first (a) an analysis was made of the eye-voice span showing the differences in the width of the span in the different grades and in the high school, and (b) the variations in the width of the span in different parts of the sentence. In both (a) and (b) the results were

[1]Buswell, G. T. *An Experimental Study of the Eye-Voice Span in Reading.* "Supplementary Educational Monographs," No. 17. Chicago: The University of Chicago, 1320. Pp. xi + 105.

grouped to show separately the characteristics of pupils with mature and immature reading habits. Following this, (c) a comparison was made to show the relationship of the width of the eye-voice span to rate of reading, number of fixations per line, and regressive movements.

In the second division (a) a very detailed analysis of the eye-voice relationship was made, showing the exact position of the eye and voice at each fixation pause, and exhibiting the variations in the width of the eye-voice span in the reading of different parts of a passage by the same individual. These results were used (b) in an attempt to explain the cause of the occasional very long fixation pauses which appear in reading records.

The third division of the investigation made use of a test device, consisting of a paragraph containing words spelled alike but pronounced differently, by which the eye-voice span in oral reading was studied in relation to the recognition of meaning in silent reading.

Method.

The apparatus and general method of the investigation were the same as used in a number of researches in the Chicago laboratory, and have been described in detail in a monograph by C. T. Gray.[1] Briefly stated, the method consisted of photographing a beam of light generated by an arc lamp, reflected first to the cornea of the eye from silvered glass mirrors, and then from the cornea through a lens to a moving film. The pencil of light changes its direction with each movement of the eye. As the subject reads a photograph is made on the moving film in the form of a sharply focused line. The shifts in this line record the movements of the eye. An electrically driven tuning fork, with a vibration rate of fifty per second, is mounted in the path of the beam of light in such a way that the light is intercepted at each vibration. These vibrations produce on the film a line of dots rather than a solid line, each dot representing a time of exactly one-fiftieth of a second. Since the film is moving continuously in the vertical plane, the record shows a vertical line of dots while the eye is fixated in a single position, and a short horizontal line when the eye is in motion in a horizontal or oblique

[1]Gray, C. T. *Types of Reading Ability as Exhibited Through Tests and Laboratory Experiments.* "Supplementary Educational Monographs," Vol. I, No. 5. Chicago: The University of Chicago, 1917. Pp. 83-90.

direction. By means of such a record the exact location of each fixation can be determined. In order to get a record of the voice the photograph was supplemented by a dictphone record of the oral reading, taken for each subject at the time the photograph was made. By means of an electrical device the dictaphone and film records were synchronized, showing the relative position of the eye and voice at different places in the reading.

Photographs taken of the readings of 54 subjects selected as follows. Two good and two poor readers were selected from each of the elementary grades above the first, on the basis of scores made in W. S. Gray's Oral Reading Paragraphs. Three good and three poor readers were selected through the co-operation of the English department from each of the four high school classes. Six adult college students were selected at random and ranked into two groups, one better in reading than the other. The entire group of subjects, therefore, included twenty-four from the elementary school, twenty-four from the high school, and six college students, each grouping being made up of equal numbers of good and poor readers.

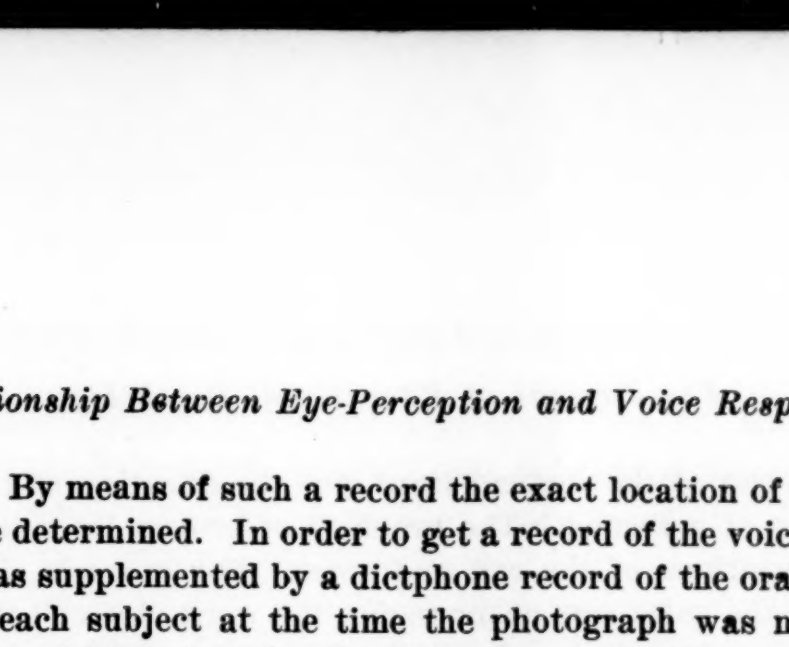
The kitten pulled at the veil and wreath
of flowers with her cunning paws. Little by
The kitten pulled at the veil and wreath

PLATE I.

First two lines show reading record of a good reader from the fourth grade. Third line shows reading record of a poor reader from the second grade.

Plate I shows sections from two typical records which will serve to illustrate the method. The short vertical lines crossing the lines of print indicate the positions of eye-fixations. The serial numbers above the verticals indicate the order of the pauses; the numbers at

the lower end of the vertical lines give the duration of the fixation in fiftieths of a second. The brackets of broken lines indicate the relative positions of the eye and voice. For example, in the first line when the voice was pronouncing the word "kitten" the eye was fixated just after the word "the." The letters "V" and "E" show the corresponding positions of the voice and eye respectively, the width of the eye-voice span at this position being 17 letter-spaces. The third line of Plate I shows the record of a second grade reader with very immature reading habits. The eye-voice span for this subject is very narrow.

Analysis of Eye-Voice Relationship.

The first division of the investigation made an analysis of the eye-voice relationship and a comparison of the eye-voice span with other factors of the reading process. The results may be summarized as follows:

1. A positive correlation was shown between a wide eye-voice span and mature reading habits. The average span for the good readers is greater than that for the poor readers in every school grade. The average width of the span for the good readers in the elementary grades is greater than that of the poor readers by 58 per cent; for the subjects from the high school it is greater by 36 per cent.

2. Width of eye-voice span has a positive correlation with rate of reading, as shown in Table I which follows.

TABLE I.

RELATION OF EYE-VOICE SPAN READING TO RATE—ALL SUBJECTS.

Number of Subjects.	Rate of Words per Second.	Average Eye-Voice Span.
1.	0-0.9	3.4
2.	1-1.9	5.7
10.	2-2.9	11.3
29.	3-3.9	12.7
12.	4-4.9	16.5

3. A negative correlation is shown between width of eye-voice span and number of fixations per line. As the span increases in width the number of fixations per line decreases.

The three results above mentioned merely confirm what one would naturally expect to find true.

4. The development of the eye-voice span does not show a consistent increase from grade to grade, but is very irregular. On the whole there is a development in width of span throughout the grades, the high school average being greater than that for the elementary grades, and the adult average greater than that for the high school. However, some good readers from the elementary grades have a greater span than many of the high school subjects. The results show that some readers develop an eye-voice span by the end of the fifth grade which is greater than the average span for all the good readers from the high school.

5. Little correlation was found between width of eye-voice span and position in the line. However, when studied in relation to position in the sentence a marked correlation was discovered. Table II shows the average eye-voice span, in number of letter spaces, for positions at the beginning, within, and at the end of sentences.

TABLE II.

AVERAGE EYE-VOICE SPAN AT THE BEGINNING, WITHIN, AND AT THE END OF SENTENCE—ALL SUBJECTS.

Subjects.	Beginning of Sentence.	Within Sentence.	End of Sentence.
Elementary subjects:			
Good readers	16.1	14.8	8.4
Poor readers	8.6	9.5	7.9
Good and poor	12.3	12.1	8.1
High-school subjects:			
Good readers	16.6	16.2	13.3
Poor readers	12.9	11.0	10.5
Good and poor	14.7	13.6	11.9
Adult subjects:			
Good readers	23.5	18.6	14.0
Poor readers	17.7	10.8	11.3
Good and poor	20.6	14.7	12.6
All subjects:			
Good readers	18.7	16.5	11.9
Poor readers	13.1	10.4	9.9
Good and poor	15.9	13.4	10.9

An examination of these data reveal two conspicuous facts: first, that the width of the eye-voice span is different at various positions in a sentence, and second, that the good and poor readers do not exhibit these differences in the same fashion. The eye-voice span is found to be relatively wide at the beginning of a sentence and relatively narrow at the end.

The fact that the eye-voice span varies with the position in the sentence is of considerable significance. If the span varied only

with the position in the line, the determining factors would be only mechanical, and would be determined by the printed form of the selection. The control of the span in that case would be a matter of the mechanics of book construction, and would be independent of any teaching factor. But if the span varies with the position in the sentence, it is evident that the content of the meaning is recognized, and that the eye-voice span is determined by thought units rather than by printed line units.

For all three classes of subjects there is agreement among the readers in that a wide eye-voice span occurs at the beginning of a sentence. The situation at the beginning of a sentence is different from that of any other position. After one has started to read, the meaning of the thought covered will carry him along to some extent, and will enable him to anticipate what is coming. At the beginning of a sentence there is no sequence of words to give one the cue to the content of the new thought. The only way to get this is to look ahead until the meaning of the sentence is partially recognized, and the kind of vocal expression needed is made clear. The good readers recognize this need for a wider span at the outset and inhibit the voice reaction until the eye has gained a considerable lead. The poor readers in the grades above the elementary school have also learned this, but evidently those in the elementary school are not mature enough in reading to recognize any special difficulty at the beginning of a sentence. Instead of making a relatively longer span, they react to the situation by a relatively shorter one. They begin to read as soon as they see the sentence, and have not learned to inhibit their reading until the eye has taken in a larger unit of meaning. This difficulty could be easily corrected by a little training in class which would teach the pupils to wait before starting to read until they get a larger unit of thought.

The evidence of all subjects agrees that there is a shorter span at the end of a sentence. The good readers have a relatively shorter span than the poor readers. The explanation of this shorter span goes back again to the fact that the sentence is the large unit of meaning. When the eye reaches the end of this unit it modifies its movements according to the meaning recognized and the voice catches up before beginning the new thought. In order that the voice shall express the thought clearly, a pause is necessary at the end of the sentence. This pause gives the eye ample opportunity for a large

eye-voice span before it is time to commence the next sentence. A poor reader pays less attention to the sentence as a unit of meaning. This is especially true of younger children who are very immature readers. For them the whole process is a more or less monotonous repetition of words as they are encountered. The eye moves along at a regular rate and the voice follows. The end of a sentence creates no special disturbance, for it is passed over with little attention. Consequently there is little change in the eye-voice span. The data for the poor readers from the elementary school would seem to indicate that some such situation exists. There is little variation in the width of the span for any position in the sentence. If the variation in the eye-voice span at the beginning and the end of a sentence makes possible a greater emphasis on meaning, the lack of such a variation may account for the fact that the subjects showing such lack are classed as poor readers.

Continuous Relationship of Eye and Voice.

A comparison of the width of the eye-voice span as measured at different positions in a paragraph showed a considerable variation from point to point in a selection read by a single subject. In order to get more complete data on the continuous relationships of the eye and voice, a very detailed analysis was made of the records of a number of subjects showing the exact relationship of the eye and voice at each word in the selection. Plate II gives a section of a typical record treated by this kind of analysis. The selection has been duplicated in parallel lines in order to show the eye-voice relationship more clearly. The upper line of the pairs may be called the eye line, and shows the position and duration of eye-fixations in the same manner as in the previous plate. The lower line of the pairs may be called the voice line. The diagonal lines connect the positions of the eye and voice for every fixation.

In the record shown in Plate II, the positions of the eye and voice were synchronized first as the voice was pronouncing the word "two" in line 1. As the voice began to pronounce "two" the eye was fixated on the last letter of the word "were," which is the fifth fixation in the line. These two points may therefore be taken as a base of measurment for the determination of the relative positions of the eye and voice at succeeding words and fixations. It will be observed by reference to the plate that 34 (20 + 14) fiftieths of a second elapsed

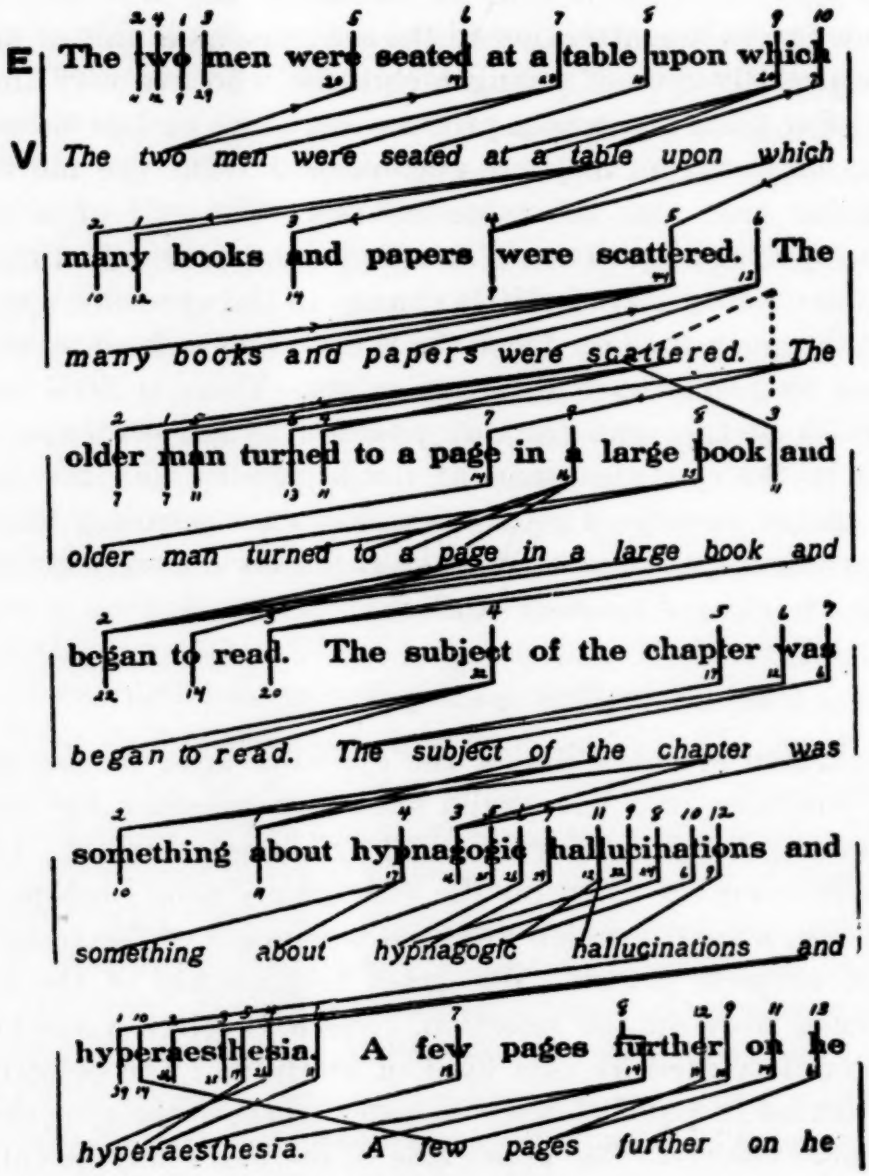
E
The two men were seated at a table upon which
V
The two men were seated at a table upon which
many books and papers were scattered. The
many books and papers were scattered. The
older man turned to a page in a large book and
older man turned to a page in a large book and
began to read. The subject of the chapter was
began to read. The subject of the chapter was
something about hypnagogic hallucinations and
something about hypnagogic hallucinations and
hyperaesthesia. A few pages further on he
hyperaesthesia. A few pages further on he

PLATE II.

Showing continuous eye-voice relationship of a good reader from the high school.

during fixations 5 and 6. The time on a stop-watch showed that at a point 34 fiftieths of a second after beginning to pronounce "two," the voice had not begun to pronounce the next word, indicating that fixations 5 and 6 were both made during the time used by the voice in reading the word "two." From fixation 5 to the end of fixation 7 a time of 57 $(20 + 14 + 23)$ fiftieths of a second elapsed. The record from the stop-watch and the dictaphone showed that during the interval from 34 fiftieths of a second to 57 fiftieths, the voice had pronounced the words "men" and "were." By a continuation of this

form of analysis the relative positions of the eye and voice were located for every word and every fixation.

Two characteristics of reading are apparent from this type of analysis. The first is the elastic nature of the eye-voice relationship. The width of the spau is evidently varied according to the demands of the content of the material. When the difficult words in the last two lines are encountered there is a very marked modifica tion of the eye-voice span.

A second type of result, which was indirectly found, is an apparent relationship between the length of fixation pauses and difficultics in reading. The average duration of a fixation pause for the subject shown in Plate II is 12 fiftieths of a second. If the three difficult words are again observed it will be seen that most of the fixations upon these words are much longer than the average. This suggested that an explanation of the occurence of long fixations might be found in the content of the reading selection.

In order to test this possibility a study was made of all fixations which were longer than 20 fiftieths of a second in the records of nineteen subjects. The readings of these nineteen subjects showed a total of 259 fixations which were over 20 fiftieths of a second in length. The selection which these subjects read contained ninety words. If the long fixations simply occurred at random there should be an average of 2.9 for each word in the paragraph. The point of interest is that the long fixations are not distributed in a random fashion with 2.9 falling on each word, but that they occur much more frequently on certain words than on others. The greatest number of long fixations fall upon the words "hyperaesthesia," "hypnagogic," and "hallucinations," these words receiving 31, 22, and 16 fixations respectively.

The interpretation of such results must be related to the development of reading habits. The records of the elementary subjects, whose reading habits are still immature, show that long fixations are frequent and occur all through the selection. Evidently long fixations are characteristic of immature types of reading. The high school and adult subjects represent more mature readers. In general they have outgrown the long fixation habits. However, when words of special difficulty or difficult phrases are encountered they return immediately to the primitive type of habits characteristic of the immature reader.

The Eye-Voice Span and the Recognition of Meaning.

The significant factor about a wide eye-voice span is that it provides a wide unit for the interpretation of meaning before the voice reaction takes place. In order to test this factor a paragraph was constructed containing a number of words which are spelled alike but may be pronounced differently. The section used is shown below, with the test words italicized.

> The boys' arrows were nearly gone so they sat down on the grass and stopped hunting. Over at the edge of the woods they saw Henry making a *bow* to a little girl who was coming down the road. She had *tears* in her dress and also *tears* in her eyes. She gave Henry a note which he brought over to the group of young hunters. *Read* to the boys it caused great excitement. After a *minute* but rapid examination of their weapons they ran down the valley. *Does* were standing at the edge of the lake making an excellent target.

The paragraph is so constructed that the meaning of the test words is not evident until the next few words in the sentence are read. For example, in line three, the word "tears" is ambiguous until the word "dress" is reached. The hypothesis in regard to the eye-voice span is that no error will be made in the reading if the span is wide enough to enable the reader to take in the word "dress" before pronouncing the word "tears." If the span is not wide enough to do this there is a strong probability of error. By varying the distance between the test words and the part of the sentence which qualifies it, a rough measure of the width of the eye-voice span can be secured. If this hypothesis is correct, subjects having a wide eye-voice span should make fewer errors than those whose span is narrower. The trial of the paragraph with a group of subjects showed the hypothesis to be correct.

Photographs of the oral reading of the test paragraph brought out two facts clearly. The first is that difficulties in the recognition of meaning are reflected in the eye movements by characteristic types of confusion. This was demonstrated by a comparison of the dictaphone and photographic records. The second fact is that subjects having a wide eye-voice span have less difficulty with such material than subjects with a narrow span.

Photographs were also taken of the silent reading of the test paragraph. The same characteristic eye-confusions occur at some of the test words as occurred in the oral records. Evidently there is in silent reading an eye-recognition span which is similar to the eye-voice span in oral reading. Comparisons of oral and silent records show that whenever difficulties with the test words are experienced, the same types of confused eye movements occur. Since subjects with a wide eye-voice span have fewer difficulties with the oral reading of such material than subjects with a narrower span, it suggests that training for a wide eye-voice span would carry over into silent reading habits.

It has been pointed out that when difficulties are encountered in oral reading the eye-voice span is immediately reduced to a primitive form. The same thing occurs in silent reading. Silent reading records showed that when the difficult words in the test paragraph were encountered, the eye returned to the word causing the difficulty just as it returned to the position of the voice in oral reading. If the difficulty in getting the meaning in silent reading is sufficiently great, there is a reversion not only to the habit of bringing the eye back to the location of the recognition of meaning, but also to the most primitive habit of silently pronouncing the words. This reinstates the most primitive form of reading where the eye, the voice, and the meaning proceed together.

The development of the reading process may be traced through three stages. First, the most primitive or immature stage of oral reading where the eye, the voice, and the meaning are all focused at the same point. Such a stage is illustrated by the reading of one of the second grade subjects where the average eye-voice span was less than the width of the average word in the selection. Second, a more mature stage of oral reading where there is a considerable span between the eye and the voice, with the center of attention near the eye, the voice reaction occurring in a semi-automatic fashion. Third, the stage of silent reading where the reader is entirely relieved of any attention to the voice and where the entire attention can be given to the eye and the meaning, making possible the development of a much higher degree of proficiency.

PROPHECY OF LEARNING PROGRESS BY BETA

GARRY C. MYERS,
Cleveland School of Education, Cleveland (Ohio) Public Schools.

During April of 1920 men of the First Recruit Educational Center, Camp Upton, N. Y., consisting of about 1400 illiterate soldiers, were given Beta. Then these men were reclassified within each grade upon the basis of the Beta ratings. Because of a quarantine in one company and considerable illness in a few other companies a few hundred men were not tested, which meant that the classification was not complete. However, after that time all men entering the Center were tested and placed into classes on the basis of their relative intelligence rating.

For the purpose of administration in that school the course was divided into six grades—not six traditional grades, but six grades for that school. The work covered by each grade as well as the requirements for promotion were highly standardized. Promotions took place every two weeks which was the minimal interval for completing the work of each grade. Because of its method of promotion this school afforded an excellent opportunity to study to what defree Beta ratings prophesy learning progress.

On the basis of all available records of men in the school June 28, 1920, a study was made of the time which had been spent in school by the men as they were found at that time in the respective sections of the several grades. For example, X with a Beta rating of 55 then in grade two, section A, had been in school a total of 31 days (including Saturday and Sunday). The problem was to find out if the low Beta men progress as rapidly as the high Beta men. Of course it were more accurate if the exact time for completing each grade were recorded for each man and if actual school days exclusive of Saturdays, Sundays and holidays, guard days, recruiting-duty days and sick days were first excluded.

Here is what was found:

Median number of days which men of each section of each grade had spent in school up to June 28, 1920.

(988 men pretty evenly distributed by sections and grades.)

Grade.	Highest Beta or Section A. Med. Days.	2nd Highest Beta or Section B. Med. Days.	3rd Highest Beta or Section C. Med. Days.	4th Highest Beta or Section D. Med. Days.
1	26	31	37	68
2	37	48	92	110
3	58	110	139	
4	110	139	129	
5	120	137	159	
6	137	155		

The median time in school by those who had completed the six grades at the Recruit Educational Center in May and June including holidays, quarantines, etc., was 173 days.

From the table one reads, for example: "On June 28, 1920, the median time which had been spent in school by Section A (section making the highest Beta rating) was 26 days as against 68 days for the lowest Beta section; the median time which had been spent by the highest Beta section of grade 2 was 37 days as against 110 days for the lowest Beta section. That the median time for the second highest Beta section of grade 4 should be higher than for the lowest Beta section of that grade is probably explained by the fact that the teacher of this section did not conform to the standards set for his section. Instead he introduced considerable experimental subject matter.

Although these figures clearly show that those ratings highest in Beta tend to progress much faster than those rating lowest in Beta, yet the figures do not tell this story well enough. In the first place, since more than fifty per cent of the men were of foreign language, many of whom did not even understand English well on entering school, some of these men relatively bright did not reveal by their Beta ratings their native capacity to learn to the same relative degree that the English speaking illiterate did—for even Beta handicapped the non-English men. This means that some men were classed in a relatively too low section. Likewise the non-English of the highest Beta section, because of English difficulties, made relatively slow progress in the first few grades regardless of their capacity. Moreover, the non-school days, sick days, quarantines, guard days, etc., which were constant for all sections of a grade,

added too many days to the high grade Beta sections in comparison with the low grade Beta sections. If, for example, an average of 10 days were deducted by these demands from the actual time in school for each section of a grade, the ratio of the time in school by the highest to that of the lowest would greatly be increased. Furthermore, the men selected from time to time for several weeks recruiting service naturally drew from the higher Beta-sections, adding thereby an undue proportion to the time-in-school-records, since there were not sufficient data available to make the proper deduction for this time. Finally, it should be added that 68 days for D Section of grade one would have increased materially if the study had been made a few months later unless a large number had been eliminated, for 67 cases of this group on June 28, 1920, had been in school a median of 165 days. The Stanford Binet ratings of these 67 cases were as follows:

Mental age in years.	No. cases.
5—5.9	5
6—6.9	15
7—7.9	24
8—8.9	19
9—9.9	4

Likewise 23 men of the lowest Beta section of the second grade had been in school a median of 217 days. Their mental age ratings were:

Mental age in years.	No. cases.
7—7.9	6
8—8.9	11
9—9.9	6

Of 19 men who had been dropped from school as "hopeless" during May and June the median time in school was 190 days with the following distribution:

Mental age in years.	No. cases.
5—5.9	1
6—6.9	6
7—7.9	7
8—8.9	3
9—9.9	2

These data seem to justify the action of the War Department in extending the Upton intelligence classification plan to all the Army's Americanization schools.

This plan, whose details are now pretty closely adhered to in all the Recruit Educational Centers, and which after its evolution at Camp Upton, the writer,* under orders, submitted to the War Department, provides that all men of a given Recruit Educational Center be examined by Beta and classified within each grade on the basis of their relative ratings, and that all men thereafter be examined only on entering the Center. It provides that the rating limits for each section be so adjusted that, although the men in the better section shall always be relatively higher in Beta rating than the men of the next lower section, it shall keep the size of the sections practically equal.

No repetition of Beta is made; but in promotion, all men of a given grade are advanced to the appropriate grade regardless of section. Then the names of those remaining in a given section together with those promoted to that section are pooled and ranked on the basis of the original Beta ratings. The number desired for the best section is then counted off from the highest of the group, those for the next section, next, and so on for the entire grade. The clinician and military psychiatrist, of course, takes care of the individual "variables."

*Then Director of Education, First Recruit Educational Center, Camp Upton, N. Y.

DEPARTMENT FOR DISCUSSION OF RESEARCH PROBLEMS

Conducted by LAURA ZIRBES

This department has a two-fold function. It aims to serve research workers as well as educators, whose work brings them in close contact with children in the schools. It hopes to accomplish this service by suggesting research studies, which will meet well-defined school needs.

In order that this service may be real and effective, the co-operation of research workers and school people is desired. Correspondence with reference to the following questions will be considered in selecting topics for future discussions.

a. Which of the studies proposed would help you to solve a practical problem?

b. What topics might well be added to this list? Replies may be addressed to: Miss Laura Zirbes, 646 Park Ave., New York City.

AN ADDITIONAL CRITERION FOR THE SELECTION OF THE ELEMENTS OF MENTAL TESTS

J. CROSBY CHAPMAN,
Yale University.

The two criteria which have been chiefly used for investigating the effectiveness of a particular element of a mental test have been

1. The increase in achievement from age group to age group;
2. The variations within each age group (coherence).

The extent to which the above criteria are fulfilled is measured by administering the particular element of the test to large groups with increasing chronological age. The first criterion obviously provides no safeguard against the training factor and the second not as much as might be desired.

In thus establishing the validity of the test elements we have been guilty of loading the dice in our own favor, we have made our task too easy. The factors which contribute to an increase in the achievement of the higher age group are—

1. Increase of intelligence;
2. Longer period of exposure to environmental and training influences. However important the second factor may be as an indication of effectiveness, it is a disturbing factor in determining the reliability of the test element as a measure of pure intelligence, defined

as the power to adapt to a novel situation. It is essential that we set up an additional criterion which will load the dice in the opposite direction. Whereas in the above situation both factors work in the same direction, we must set up a criterion in which they work in opposite directions.

That such a criterion is possible the following simplified case will illustrate. Suppose instead of merely trying out the test element on large groups of increasing mental age, in addition each element of the test was submitted to a group of chronologically young, but mentally bright children, and to a second group of children, mentally dull but chronologically old. To make the situation more specific, let us suppose that each test element is tried out on

1. Fifty children of age 8 who show on any composite test a mental age of 10.

2. Fifty children of age 12 who show on any composite test a mental age of 10.

We should then establish as our additional criterion of the excellence of a particular test element the extent to which the younger group exceeded the older group.* In this case the environmental factor would be working against the intelligence factor. In other words we would have loaded the dice against ourselves. Could we not in this manner arrange our tests in ascending order of merit, according to the way in which they differentiated more and more successfully in favor of the younger group? Would it not happen that the boasted merits of the arithmetic tests and vocabulary tests would very much dwindle, possibly melt into thin air, in favor of those tests which less obviously are subject to environmental factors? In this way we might conceivably eliminate certain tests which measured by present criteria give good indications of differentiating power but which under more refined testing would show their limitations.

A more rapid method of attacking the problem would be to compare the scores obtained in any of the recognized composite group tests by the young-bright, and the old-dull children. We could take the chronological ages as far as possible apart, consistent with securing a reasonable number of each of the groups scoring approximately the same number of points. We could then investigate the extent

*See the work of Chatzen on a less exacting criterion.

to which these two different groups had secured their totals on the same elements of the test.

The author is at present collecting and reviewing experimental data which will shortly be published with a view to testing the extent to which the application of this "criterion of environmental influence" will change the present estimation of the relative effectiveness of the various test elements. The publication of this note has been hastened by the consideration that many individuals have the necessary data, in greater quantity than the author, with which to test the fruitfulness of this exacting criterion.

Unless the above criterion is applied we get the anomalous situation which at present exists, that the mental ages of children in a school system are a function of the enlightenment of that school system with reference to retardation and acceleration! It is obviously absurd to prejudice the mental rating of an extremely bright child simply because a cast iron school system has, through ignorance, denied the promotions necessary for a certain facility in, let us say, the arithmetic and vocabulary tests which contribute to high mental standing! The eventual outcome of the changes in the mental test elements must be to prove that the brighter children are really more intelligent than the present tests indicate, and that the duller children are even less capable than present measurements reveal.

HOW CAN GROUP TESTS BE BETTER ADAPTED TO VERY YOUNG CHILDREN?

Dr. Agnes Rogers remarked in a recent lecture at Teachers College that some makers of group tests for children of kindergarten age had made some very ingenious provisions for securing attention and enlisting interests. She also believes that the consideration of the following proposals in determining the content or form of such tests might add materially to the successful administration of group tests to young children:

1. That when pictures are used, their size and clearness be determined psychologically. 2. That the content take more careful account of the child's range of experience. 3. That the tests be divided into parts, printed on separate small sheets or cards, to

avoid the confusion attending the presentation of a large number of pictures or an unwieldly leaflet. 4. That there be a standardized requirement for giving separate parts of the tests at intervals rather than in immediate succession. 5. That there be a standardized period for practice in the use of paper and pencil as a test preliminary for children of kindergarten age especially. 6. That there be some method to reduce the tendency to communicate, which is especially difficult to control in absurdities tests.

Dr. Rogers thinks that the group test is particularly adapted to the measurement of the timid child, especially when individual tests are administered by strangers.

The present writer suggests that there be some group activity immediately preceding the test, so that pupils unaccustomed to group responses are not unduly hampered by the requirements of group testing. This could be in the nature of a prescribed game in which pupils follow instructions without conversation or individual variation. Such group activity would be part of the standardized test instructions and lead to more uniform responses. The objective evaluation of the curriculum and methods used with young children will no doubt be facilitated by the use of tests which take careful account of the limitations of pupil material. L. Z.

Dr. Ernest Horn of the University of Iowa and Dr. William McCall of Teachers College, Columbia, will contribute discussions to this department in the next issue.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

1. A new monograph on the psychology of reading. Dr. Buswell's study[1] of the eye-voice span is the fifth of The University of Chicago monographs which deal exclusively with reading. The laboratory results which it reports throw much light on problems suggested in studies by Quantz, and Judd, and C. T. Gray. With the apparatus used by Gray the author obtained photographic records of the eye-movements of fifty-four subjects, selected to include equal numbers of good and poor readers, from the second grade through first year college. By the simultaneous use of the camera, dictaphone, and ingenious synchronizing devices, the records were made to reveal the eye-voice span, or the point at which the eye was fixated when the voice pronounced any given word in the selection.

The study was very carefully set up to reveal any significant relationships which might exist between eye-voice span and other elements of reading ability. Plates, figures, and tables are presented so that the data may be consulted by the reader and conclusions can be studied by reference to the data on which they are based. There is cumulative evidence that the element studied has significant relations to comprehension, rate, and the quality of reading, as well as to number and length of fixations. The data lead to the conclusion that eye-voice span varies for the same reader within a given selection for causes which have to do with content and meaning, rather than with position in the line. Mature readers have longer spans, which give them a distinct advantage in the recognition of meaning and the avoidance of error, as demonstrated by tests containing words difficult to understand or interpret. When difficulties cannot readily be negotiated, mature readers revert to the primitive type of habits characteristic of the immature reader.

The study shows that the eye behaves similarly when encountering difficulties in silent reading. The immature reader evidently derives

---

[1]Buswell, G. T. *An Experimental Study of the Eye-Voice Span in Reading.* Supplementary Educational Monographs, Vol. III, No. 17, Chicago. University of Chicago, Dept. of Education, 1920. Pp. XI + 105. $1.00. Paper.

less benefit from silent reading as his eye and voice proceed more nearly together. There is hence no opportunity to adjust between seeing and saying. If this is a measure of immaturity, Dr. Buswell's study leads to the conclusion that the advantage of a wide attention span should be made available by special training during the early grades. The laboratory studies show that good readers in the lower grades have the eye-habits of mature readers. The extent to which a wide span may be developed, or to which it is dependent on native capacity, is a problem well worth careful investigation.

L. Z.

---

2. A New Handbook on Sex Education. While Dr. Galloway's new book[2] is intended primarily for the use of college men there is much in it which should give it an audience in other circles. The author feels that the college man will be much more likely to develop a satisfying point of view in matters pertaining to sex if he considers his responsibility to younger boys and prepares himself for leadership in the character development of others. The organization of the book is noteworthy. Discussion topics are listed at the beginning of each chapter. An introductory statement is then followed by brief specific questions and answers, with references to source materials.

The introductory statements taken together are a sane and clear statement of the issues, replete with well chosen illustrations which cannot fail to impress the careful reader with the fundamental philosophy upon which the author builds his sex-education program. The dynamic character of appetites and desires must be considered in any thoroughgoing plan for character development. Our reactions to appetites and impulses are measures of our control and character, and are a means of self-development which cannot be sidestepped or eliminated without grave consequences.

Only ignorance of the social and ethical potentiality of refined and redirected appetites and desires leads to a social and educational program which capitalizes the dynamic human factors underlying behavior. Dr. Galloway's book makes its greatest contribution by clarifying the issue, and showing its relationship to individual and social progress.

L. Z.

---

[2]Galloway, T. W. *The Sex Factor in Human Life.* The American Social Hygiene Association, New York City, 1921. Pp. 142.

3. A report of the use of mental tests in school administration. The aim of this study[3] is stated very simply: to show how the results of certain educational and psychological tests may be used as aids in problems of school administration. The results reported were obtained from tests of 125 boys and girls of the Junior Division of the University High School, Eugene, Oregon. These pupils were drawn from two types of homes:—some came from families who were professionally connected with the University; another group was drawn from the homes of unskilled workmen.

All the pupils in the school were tested with the Stanford Revision of the Binet Scale, the Army Group Examination Alpha, and the Chicago Group Intelligence Test, devised by Freeman and Rugg. The results for the total group, stated in terms of the median, are as follows: Mental Age, 14-9; Intelligence Quotient, 107.0; Alpha, 89.0; Chicago, 38.0. These results show clearly that the children of this group are a selected group, with intelligence above that of the average for this age. When the results of each group test are correlated with the Binet scores, the correlation for Alpha is found to be 0.73, in comparison with 0.62 for the Chicago Scale. The author thinks this difference may be accounted for by the fact that the Chicago scale requires 15 minutes less time than does the Alpha test.

Six experienced teachers, after having received some instructions in methods of making intelligence ratings, rated all the pupils they each knew. These estimates of intelligence were correlated with the I.Q.'s. The resulting coefficients ranged from 0.61 to 0.72, with an average of 0.68. The pooled ratings also gave an r of 0.68. This high correlation is contrasted with a lower one of 0.50, which results from the correlation of the estimates of ten inexperienced practice-teachers, with the I.Q.'s.

The question of acceleration and retardation is discussed, and it is shown that, on the basis of Mental Age, 31.1 per cent are accelerated, 22.7 per cent are normal and 46.2 per cent are retarded. Of 36 pupils who are retarded by usual (actual age) standards, 41.7 per cent are accelerated, by Mental Age. Of 40 who are accelerated by the usual standards, 55.0 per cent are retarded, by Mental Age. The author comes to the conclusion that "the accelerated pupils upon the basis of actual age tend to become retarded upon the basis of

---

[3]Ruch, Giles Murrel. *A Study of the Mental, Pedagogical and Physical Development of the Pupils of the Junior Division of the University High School, Eugene, Oregon.* University of Oregon, Eugene, Oregon.

mental age, and that the reverse holds for the group considered retarded by the usual standards"—a conclusion which other studies corroborate.

Seven tests of educational achievement were applied to the group:—Courtis Standard Research Tests in Arithmetic, Series B; Stone Reasoning Tests; Gregory Language; Ayres Spelling; Handwriting scored by the Ayres Scale; Kansas Silent Reading; Douglas Tests for elementary Algebra. The results for grades VII and VIII are compared with standard scores; they prove to be "well above the norms in all tests except in the fundamentals of arithmetic as shown by the Courtis Tests, and in writing." A study of the Stone Reasoning Tests seems to indicate that the low scores in computation are due to poor training in the lower grades of the system. The poor showing in handwriting is explained by the fact that too little time on the school program is allotted to penmanship practice.

Certain anthropometric measurements were obtained to supplement the above results. In height and weight no significant deviations from the available norms were found. When compared with Smedley's results the Eugene boys and girls show "markedly greater development of lung capacity at all ages." From certain correlations obtained, the author concludes that there "exists a positive relationship between the vital index and mental ability. This relationship is not very perfect and is probably of indirect value."

The most important of the conclusions based upon the results just discussed are as follows:

1. While it is true that trained teachers of long experience can estimate intelligence with a high degree of accuracy, they do occasionally make serious errors. It is in these exceptional cases that intelligence tests are most needed. "The fact that comparisons of the ratings obtained by the two methods show discrepancies leads to deeper analysis of specific abilities by teachers, and to a truer understanding of the complex relationships between the many factors combining to determine school achievement."

2. The results of such tests may be of great assistance in helping to determine the speed of school advancement of each individual. If forced promotions are necessary, the results of the tests will aid in selecting the pupils for such promotion. Especial provision, in separate classes, should be made for children of high native ability.

3. Such tests are of value in determining the *special* needs of each child.

4. "Mental deficiency is a far more powerful factor in the cause of retardation than has been commonly supposed."

5. "Correlation of abilities in school subjects with general intelligence shows the highest coefficients in the language and reasoning tests, and the lowest in spelling and writing. School marks and arithmetical ability are intermediate."

6. Pupils who are to be classified as of average ability are "somewhat more likely to realize results in school according to their ability than are either very able or very dull pupils."

7. "In view of the value of the results of intelligence testing from the administrative angle the amount of time required for tests cannot reasonably be held to be prohibitive." G. L. Coy.

Ohio State University.

---

4. A survey of material used for memory work in city public schools. This bulletin[4] attempts to present the present day American city public school practice in regard to the specific material used for memory work and the amount of such work required. There are given in detail the method of procedure and the results of a survey of fifty city courses of study, selected from a much larger group, primarily on the basis of complete lists of memory material offered and the specificness of the memory work requirements. Due to indefiniteness in the courses of study which were surveyed, only the data which refer to poetry memory work are at all complete.

From the complete lists of 2,435 poems mentioned in the various courses of study, the investigator selects the 329 that were mentioned in five or more courses. These selected poems are presented in various suggestive and useful tabular forms. The first of these tables gives the titles and the poems arranged in the order of frequency of mention and shows, as well, the number of different cities mentioning, and a weighted value for each poem. This value is obtained by using the factors 1, 2, 3, and 4 respectively for mentions in which the poem is given "for study"; is suggested for memory, other poems being required for memory; "suggested for memory without distinction"; "required for memory." A second table shows

---

[4]Bamesberger, Velda C. *Standard Requirements for Memorizing Literary Material.* University of Illinois, Bureau of Educational Research. Bulletin No. 3. Pp. 93.

the distribution of the mentions of each poem according to the school grade for which it is recommended by the various courses of study. The results of the final compilations of the data in this form are then so presented as to show just which poems are now tending to be used for memory work in each grade from the first to the eighth inclusive. A group of "Preferred List of Poems for Memory Work" arranged by grades and consisting of 112 poems presents, we are inclined to agree with author, "the best short lists which an investigation of present courses of study can yield." These preferred lists represent in the main the agreements in graded classifications of poems as ascertained from the data of this investigation and those of a similar investigation made by Lewis Atherton and published in 1914.

These lists should prove suggestive and helpful to all teachers. To those teachers who have not sufficient literary training, or initiative, or time left them by the frequently overburdensome schedules, these lists may well serve as excellent guides to keep them within the limits set by the accepted standards for this type of work. Those teachers who wish to make use of the lists, as well as the educationalists who may desire to do experimental work upon the relative values of the various poems listed, will discover that much aid will be given them by the excellent arrangement of the poems into a "Finding List" arranging the poems by author and giving titles, first lines, suggested grade, and text references.

We feel very strongly, however, that every teacher making such use of these lists should be at the same time alive to the inadequacies inherent in lists thus compiled. The author is, herself, well aware of the practical limitations of the lists. She emphasizes the fact that the lists are "based on agreement between courses of study. If there is any constant bias—any convention which has grown up whereby one writer of a course imitates others—these results will likewise be biased in the same direction. If, for example, there is a tendency to assign to elementary school children poems which they cannot understand or appreciate, this tendency will be evident in the resultant lists."

Certain tendencies, which might well be considered as biased tendencies, are clearly shown by the author to influence the composition of the lists. There is the tendency to unduly favor American at the expense of English authors, shown not only by the many more

poems chosen from American authors, but also by the inclusion of work from a large number of minor American poets. The work of foreign language writers is practically excluded, as is also the work of recent poets.

In the editorial introduction is found the suggestion "that any desirable form must be based upon present practice as the point of departure." Perhaps, after all, the chief value of this present work may be found to consist in supplying this necessary point of departure.

HARRY W. CRANE.

Ohio State University.

---

5. An Experimental Study of Memory. Using a variety of materials (words, geometrical forms, proverbs, and syllables), with approximately 100 adults, and 600 school children, as subjects, Dr. Achilles has made an intensity study of Recall and Recognition.[5] As in earlier investigations, it is found that the number of items recognized surpasses the number recalled by amounts depending upon the material and subjects. With relatively homogeneous groups, a low but positive correlation between the two processes is found. Other results conform interestingly to recent theories that mental behavior depends upon many, more or less specific, rather than a few general capacities. These correlations between Recall for different types of material are low (averaging around 0.10 with large P. E.'s). The correlations for Recognition of different materials are similarly low but positive. Women and girls are in general found to be slightly superior to men and boys. Both Recall and Recognition increase rather uniformly with age (8.5-16.5 years) and with grades (4-8). The younger pupils in a grade usually surpass the older. The tests were applied to insane patients but produced results of no diagnostic significance. A final chapter gives the details of the Recognition process, showing in general that the subject is more correct in judging a thing *as not seen* before than as *seen* before.

A. I. G.

---

[5]Achilles, Edith Mulhall. *Experimental Studies in Recall and Recognition.* Archives of Psychology, 1920. No. 44, pp. V+80.

JUN 1-